

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 22:13:32 ; Search time 156.617 Seconds
(without alignments)
7896.639 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229

Sequence: 1 caggagagccaccagccctca.....aaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	48.2	2.2	8607	4 US-10-204-708-72	Sequence 72, Appl
2	47.4	2.1	1937	4 US-09-647-143-1	Sequence 1, Appl
3	45.8	2.1	1048	4 US-09-489-847-38	Sequence 38, Appl
C 4	45.4	2.0	615	4 US-09-252-991A-10075	Sequence 10075, A
5	45.4	2.0	1126	4 US-09-252-991A-9937	Sequence 9937, Ap
6	45.4	2.0	1926	4 US-09-252-991A-9758	Sequence 9758, Ap
C 7	43.2	1.9	2823	4 US-09-252-991A-7100	Sequence 7100, Ap
8	43.2	1.9	3456	4 US-09-252-991A-7025	Sequence 7025, Ap
9	42.6	1.9	558	4 US-09-252-991A-6738	Sequence 6738, Ap
10	42.6	1.9	787	4 US-09-621-976-1878	Sequence 1878, Ap
11	42.6	1.9	2346	4 US-09-252-991A-6782	Sequence 6782, Ap
12	42.4	1.9	630	4 US-09-252-991A-3386	Sequence 3386, Ap
13	42.4	1.9	1647	4 US-09-252-991A-3344	Sequence 3344, Ap
C 14	42.2	1.9	1089	4 US-09-796-766-13	Sequence 13, Appl
15	42	1.9	804	4 US-09-252-991A-14038	Sequence 14038, A
C 16	42	1.9	807	4 US-09-252-991A-14227	Sequence 14227, A
17	42	1.9	2083	4 US-09-716-129-41	Sequence 12941, A
18	41.8	1.9	159	4 US-09-621-976-17448	Sequence 17448, A
C 19	41.8	1.9	1653	4 US-09-252-991A-10546	Sequence 10546, A
20	41.8	1.9	1701	4 US-09-252-991A-10316	Sequence 10316, A
21	41.6	1.9	1212	4 US-09-182-145-34	Sequence 34, Appl
C 22	41.6	1.9	1212	4 US-09-182-145-35	Sequence 35, Appl
23	41.4	1.9	276	4 US-09-621-976-16611	Sequence 16611, A
24	41.4	1.9	2394	4 US-09-800-729-33	Sequence 33, Appl
25	41.2	1.8	244	4 US-09-621-976-484	Sequence 484, Ap
26	41.2	1.8	298	4 US-09-621-976-3871	Sequence 3871, Ap
27	41.2	1.8	1161	1 US-08-086-439C-2	Sequence 2, Appl

28	41.2	1.8	1161	1 US-08-434-877-2	Sequence 2, Appl
29	41.2	1.8	1367	3 US-08-475-742-3	Sequence 3, Appl
30	41.2	1.8	1367	4 US-08-261-293-3	Sequence 3, Appl
31	41.2	1.8	1370	1 US-08-056-051-1	Sequence 1, Appl
32	41.2	1.8	1370	1 US-07-928-611-17	Sequence 17, Appl
33	41.2	1.8	1370	2 US-08-487-811A-17	Sequence 17, Appl
34	41.2	1.8	1370	3 US-09-060-694-17	Sequence 17, Appl
35	41.2	1.8	1370	4 US-09-378-074-17	Sequence 17, Appl
36	41.2	1.8	1370	5 PCT-US93-07370-17	Sequence 17, Appl
37	41.2	1.8	1466	1 US-08-056-051-3	Sequence 3, Appl
38	41.2	1.8	1466	1 US-07-928-611-19	Sequence 19, Appl
39	41.2	1.8	1466	2 US-08-487-811A-19	Sequence 19, Appl
40	41.2	1.8	1466	3 US-09-060-694-19	Sequence 19, Appl
41	41.2	1.8	1466	4 US-09-378-074-19	Sequence 19, Appl
42	41.2	1.8	1466	5 PCT-US93-07370-19	Sequence 19, Appl
43	41.2	1.8	1504	4 US-09-016-434-1276	Sequence 1276, Ap
44	41.2	1.8	1507	4 US-09-453-323-1	Sequence 1, Appl
45	41.2	1.8	1610	1 US-08-056-051-5	Sequence 5, Appl

ALIGNMENTS

```
RESULT 1
US-10-204-708-72/c
; Sequence 72, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBERCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 72
; LENGTH: 8607
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-72
```

Query Match 2.2%; Score 48.2; DB 4; Length 8607;
Best Local Similarity 54.1%; Pred. No. 0.022;
Matches 98; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY	2046	TTCAATAAAGCTGTAAGCTGACGACATTTGAAACATTTAAGCTGAT	2105
DB	4244	TTCAATTAATCACTTTTAAAAAACTTAATCTTTAAAAAAATTAATTAATCTTCAAT	4185
QY	2106	TCAGAGTCGGGAACCTTAGTCTTCTGAATTCACAGACGCCACCTTAGTATACGC	2165
DB	4184	TCATACCTACTCAAAAAATCTTAATTCACAAATTAACCAAAAAATTAATTAATTA	4125
QY	2166	CCAATCTAATGCTTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	2225
DB	4124	ACCCCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	4065
QY	2226	A 2226	

Db 4064 A 4064

RESULT 2

US-09-647-143-1
; Sequence 1, Application US/09647143
; Patent No. 6680196
; GENERAL INFORMATION:
; APPLICANT: Batria, Surinder K.
; APPLICANT: Hollingsworth, Michael A.
; APPLICANT: University of Nebraska Board of Regents
; TITLE OF INVENTION: No. 6680196 Gene That is Amplified and
; TITLE OF INVENTION: Overexpressed in Cancer and Methods of Use Thereof
; FILE REFERENCE: UNMC63121
; CURRENT APPLICATION NUMBER: US/09/647,143
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US99/06633
; PRIOR FILING DATE: 1999-03-26
; PRIOR APPLICATION NUMBER: 60/079,649
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1937
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-647-143-1

Query Match 2.1%; Score 47.4; DB 4; Length 1937;
Best Local Similarity 83.1%; Pred. No. 0.014; Indels 0; Gaps 0;
Matches 54; Conservative 0; Mismatches 11;

QY 2165 CCCAACTAATGATTTAATTAATCAATATCTGTTAAAAA 2224
Db 1854 CCCAACTTGTCTGTTAATTAAGTCAATCTCTTTAAAAA 1913
QY 2225 AAAAA 2229
Db 1914 AAAAA 1918

RESULT 3

US-09-489-847-38
; Sequence 38, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 38
; LENGTH: 1048
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-847-38

Query Match 2.1%; Score 45.8; DB 4; Length 1048;
Best Local Similarity 56.2%; Pred. No. 0.027;

Matches 86; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 2077 ATTGAAACATTAACTGAGCTTGATTCAGAGTGGGAACCTTATGTTCTATCTGAA 2136
Db 804 ACTGTCACTACTACTATCTGATCTGCGACAGGATTCAGAAATCTGGAACCTTA 863
QY 2137 TCCAGACGACCACTTATGATCTGCCCCAACTAATGATTAATTAATCAAAATAC 2196
Db 864 TTACGTGTGCCAAATATTACTCTCTTATGATCAACAGTCAATTAATATCTTA 923
QY 2197 TCGTTAAAAA 2229
Db 924 TTAAGAAAAA 956

RESULT 4

US-09-252-991A-10075/c
; Sequence 10075, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 10075
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-10075

Query Match 2.0%; Score 45.4; DB 4; Length 615;
Best Local Similarity 48.6%; Pred. No. 0.025; Indels 0; Gaps 0;
Matches 124; Conservative 0; Mismatches 131;

QY 318 TACATCCAGCTCATCCCATTTCTTCTGCTGGGCTTTAGTGAAGCTGTGATACC 377
Db 450 TTCAATGCTGCTATGCTGTTCGCGCTGTGCGCCCTGCGGCCCTTACGATCCG 391
QY 378 CGCTGTGAACCACTAGTGAAGAACTGCGCGCCGACCGCTCATGAGCTGATCG 437
Db 390 AGCGAGCAAGTCCGGAACTTCTGCTGAACCCCGCATCTGCTGGAAGGGGCAAGCG 331
QY 438 GCGTCTGTCGAAGCAAGGACGAGCAAGCCGCGCTGTCGGCGACCGCTATCCGTAC 497
Db 330 CGCTTCTCTGCTGCGACGACGAGCTGGGCGCGACCTGCTTCGCGGCTGATCCAGCG 271
QY 498 GCCAACCTGGGCAACGTGCTCATCTGCGGACGCTCAAGCCGAGTCAAGGCGCTTC 557
Db 270 GCCCGCTGTGCTGTGATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 211
QY 558 CCCAGGCCCGACGAC 572
Db 210 CTCCTGGGCTGCTC 196

RESULT 5

US-09-252-991A-9937
; Sequence 9937, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 9937
LENGTH: 1116
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9937

Query Match 2.0%; Score 45.4; DB 4; Length 1116;
Best Local Similarity 48.6%; Pred. No. 0.035;
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 318 TACATCAGCTCATCCCATTTCTTCTGCTGCTGAGCTTTTACGTGACGCTGCTGAC 377
DB 263 TTCTATGCTGTGATCGTGTCTGCGCGCTGTGCCCCCTGGGTGCGCCCTACGATCCG 322
QY 378 CGCTGTGGAACAGTACGAGAACCTGCGGCGCCGACCGCTCATGAGCTGATGTCG 437
DB 323 AGGAGACAGTTCCGGGACTTCTGCTGACCCCGCATCTGCTGGAAGCGGCGACGCG 382
QY 438 GAGCTTCGTCAGAGCAAGAGCAAGAGCGGCTGCTGCGGCGCAAGCTCATCCGCTAC 497
DB 383 CGCTTCCTGCTCGGACCGACGAGCTGGGCGGACCTGCTCGGCGCTGATCCAGCGC 442
QY 498 GCCAAGCTGGGCAACGCTGCTCATCTGCGGACGCTCAACCGGATCTCAAGCGCTTC 557
DB 443 GCCCGCTGTGCTGCTGATCGGCTGTCTCTCGGTGTGATCTCGGATCCCGGCGATC 502
QY 558 CCCAGGCGCCAGCAGC 572
DB 503 CTCCTCGGCTGCTC 517

RESULT 6
US-09-252-991A-9758
Sequence 9758, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 9758
LENGTH: 1926
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-9758

Query Match 2.0%; Score 45.4; DB 4; Length 1926;
Best Local Similarity 48.6%; Pred. No. 0.049;
Matches 124; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 318 TACATCAGCTCATCCCATTTCTTCTGCTGCTGAGCTTTTACGTGACGCTGCTGAC 377
DB 1129 TTCTATGCTGTGATCGTGTCTGCGCGCTGTGCCCCCTGGGTGCGCCCTACGATCCG 1188
QY 378 CGCTGTGGAACAGTACGAGAACCTGCGGCGCCGACCGCTCATGAGCTGATGTCG 437
DB 1189 AGGAGACAGTTCCGGGACTTCTGCTGACCCCGCATCTGCTGGAAGCGGCGACGCG 1248
QY 438 GAGCTTCGTCAGAGCAAGAGCAAGAGCGGCTGCTGCGGCGCAAGCTCATCCGCTAC 497
DB 1249 CGCTTCCTGCTCGGACCGACGAGCTGGGCGGACCTGTCTCGGCGGTGATCCAGCGG 1308

QY 498 GCCAAGCTGGGCAACGCTGCTCATCTGCGGAGCTGACGACCGGAGTCTACAGCGCTTC 557
DB 1309 GCCCGCTGTGCTGCTGATGAGCTGTCTCTCGGTGATCTCGGATCCCGGCGATC 1368
QY 558 CCCAGGCGCCAGCAGC 572
DB 1369 CTCCTCGGCTGCTC 1383

RESULT 7
US-09-252-991A-7100/c
Sequence 7100, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7100
LENGTH: 2823
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7100

Query Match 1.9%; Score 43.2; DB 4; Length 2823;
Best Local Similarity 52.8%; Pred. No. 0.24;
Matches 93; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 414 GACCGCTCATGAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 473
DB 2478 GACCGCTCATGAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2419
QY 474 CTGCGGCGGACGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 533
DB 2418 TACCACTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2359
QY 534 AGACCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 589
DB 2358 AGCGGAGCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2303

RESULT 8
US-09-252-991A-7025
Sequence 7025, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 7025
LENGTH: 3456
TYPE: DNA
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7025

Query Match 1.9%; Score 43.2; DB 4; Length 3456;
Best Local Similarity 52.8%; Pred. No. 0.27;

```

RESULT 9
US-09-252-991A-6738
; Sequence 6738, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6738
; LENGTH: 558
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6738

```

Query Match	1.9%	Score 42.6	DB 4	Length 558
Best Local Similarity	46.7%	Pred. No. 0.13		
Matches 135	Conservative 0	Mismatches 154	Indels 0	Gaps 0
QY	281	GCTGATGTTTGAGAAACTGACTCTGTATTTGCGACAGCTACATCCAGCTCATCCCATTTTC	340	
Db	16	GCCGCTACCCGAAAAAACGACACGAGGAGGCGCTGAGCGACGCGCGCTACTGTGCCATTCG	75	
QY	341	CTTTCGTCTGGGCGTTTCTTACGTGACCGGTGTGTGACCCCGCTGGTGGACCACTGATGAGAA	400	
Db	76	GACCGCTGCGGAAGCTGACTGACCTTGAGCGTGGCGCGCTGAACCTTGGCGCGGTGACCTTCGAGAA	135	
QY	401	CTTCGCGGTGGCCCGACCCGCTTCATGAGCGCTGTGTTCGGGCTTTCGTGTGAAAGGACAGGACGA	460	
Db	136	GCTGCGGATGTGACGACGCGCCAGCTTAAGCTTCGCGGCGGCGGCGGTATATCAAGCTTCGA	195	
QY	461	GCAAGGCGCGGCTGTCTGCGCGGACGCTTATCCGTTACGCCAACCTGGAGCAAGTGCCTCAT	520	
Db	196	CGACATGCGCGGCGGAGCTGTACCAACGCGCGCTTCAACGCGCAAGGCAACCTGAGTGTACG	255	
QY	521	CTTCGCGCAAGCTTCAGACACCGGAGGCTTCAAGCGCTTCCCAAGGCGCCAG	589	
Db	256	CCAGGACAGCGCGCTGTCTGACCGCCACCAACCACTGCGCCAAAGTGGCG	304	

```

RESULT 10
US-09-621-976-1878
; Sequence 1878, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976

```

```

Query Match      1.9%; Score 42.6; DB 4; Length 787;
Best Local Similarity 67.4%; Pred. No. 0.16;
Matches 60; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

Qy      2141 AGGACGACACCTTATGATATCTGCCCAACTATATGATTTATTAATCAATATCTCGT 2200
Db      657  AGCGACGACGTATATGTAACGGGGCATCTCTCCATTTGCATGCAAAAAAAAAAAAAA 716
Qy      2201 TAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2229
Db      717  AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 745

```

```

RESULT 11
US-09-252-991A-6782
; Sequence 6782, Application US/0925291A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 6782
; LENGTH: 2346
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-6782

```

Query Match	1.9%;	Score 42.6;	DB 4;	Length 2346;
Best Local Similarity	46.7%;	Pred. No. 0.31;		
Matches	135;	Conservative	0;	Mismatches 154; Indels 0; Gaps 0;
QY	281	GCTGATGTTTGAGAAACTGACTCTGTATATGAGACAGTACATCCAGTCATCCCATTTTC	340	
Db	1407	GCCGCTACCGGAAAAACCGACCAAGCAGGCGCTGAGAGAGAGCGCCGCTTACTGCGCATCGC	1466	
QY	341	CTTCGAGCTGGGGCTTCTACGTGAGACGCTGTGTGTGACCCCGTGTGAAACCAATACAGAA	400	
Db	1467	GACCTCCGCAAGCTTGACCTTGACCTTGACCTGAGCGCTGAACTTTCGCGCGGCTGACCGTGAAGA	1528	
QY	401	CCTGCCGTGGCCCGACCGCCCTCATGAGCCCTGTGTGTGGGCTTTCGTGAAAGCAAGAGA	460	
Db	1527	GCCTGCGATTCGACGACGCGCAGACCTGGAAGCTGGCGGCGCGAGGCGCGGTATATACGCTTGA	1586	
QY	461	GCAAGCGCGGAGCTGTGGGGGCGACGCTCATCCGCTTACCGCAACTGGGCAAGCTGTTCAT	520	
Db	1587	CGACATGCGCGCGCGGCTGTATCAACGCGCGCTTCAACGCAAGGCAACCTTCGATGTATACG	1646	
QY	521	CCTGGCGAGCGTCAAGCACCGCGAGTCTTACAAAGCGCTTCCCGACGCGCCAG	569	

Db 1647 CCAGGACAGCGCGTGTGACCGCCACCAACACATCGCCCAAGTGC 1695

RESULT 12
US-09-252-991A-3386
Sequence 3386, Application US/09252991A

```

/ GENERAL INFORMATION: Rubenfield et al.
/ APPLICANT: Marc J.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 3142
/ SEQ ID NO 3386
/ LENGTH: 630
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
/ US-09-252-991A-3386

```

Query Match	1.9%	Score 42.4	DB 4	Length 630
Best Local Similarity	49.1%	Pred. No. 0.16		
Matches 112	Conservative 0	Mismatches 116	Indels 0	Gaps 0

Qy	309	TGGGACAGCTAACATCCAGGTATATCCCAATTTCTTGAGCTGGGGCTTCTAAGTAAAGCTG	366
Db	352	TGGGACGTCGCCGACGAAACGATTCGAAGCCTCTGGCCGAGCTGTGTGACGGGGCTGTCTTC	411
Qy	369	GTCGTACCCCGCTGTGTGAACCAAGTACGAAGAACTGCGCGTGGCCGACCGGCTCATAGC	428
Db	412	GTCAGCCATTGTGTACCGGGCGCCCGGGCCATCTGCGCGCAGTGGCGGTTCAACTGTTGGCC	471
Qy	429	CTGTGTCTCGGGCTTGCTCGAAGGCAAGGACGACAGGCGCGACTGTGCGGCGCAAGCTC	488
Db	472	ATGTGTTCAACGGCGCGACGAAACGGAGATTCGAAGGCTATCCGACAGATATCCGCTCTG	531
Qy	489	ATCCGGTAAAGCCAACTCTGGGCAAGGTGCTCATCTCGGCGCAAGCTGTACG	536
Db	532	CTCGGGAGAGACTGTGCGCGCCGACAGATGTCTGTGAACCAACCCGCAATC	579

RESULT 13
US-09-252-991A-3344
Sequence 3344, Application US/09252991A

```

/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 3344
/ LENGTH: 1647
/ TYPE: DNA
/ ORGANISM: Pseudomonas aeruginosa
/ US-09-252-991A-3344

```

Query Match	1.9%	Score	42.4	DB	4	Length	1647
Best Local Similarity	49.1%	Pred. No.	0.29				
Matches 112; Conservative	0	Mismatches	116	Indels	0	Gaps	0

QY	309	TGGACAGTATCATCCAGTATATCCCATTTCTTTCGGTCCGGGCTTCTACAGGAGCTG	368
Db	212	TGGACAGTCCCGACGAAACGATCGAACGCTCGCGCGGCTGTGTCGGCGGGCTCTTC	271
QY	369	GTCGTGACCCCGCTGTGTGAAACCATGACGAACTGTCCGTGGCCCGACCGCTCATGAGC	428
Db	272	GTCAGGCATTGTCTACCGGGCCCGGCGCATCTGCACGATGTGGCGCTCAACCTGTCGCC	331
QY	429	CTGTGTTCGGGCTTTCGTGTGAAAGCAAGACGACGAGCGCGCTGTTCGGCGGACGCTC	488
Db	332	ATGTGTTCAACGGCGGACGAAAGCGGAGATCGAAGGCTATCGCCACGAGATCCGCTGCTG	391
QY	489	ATTCGGTACGCCAACCTGGGCAACGCTGTCAATCCGTGACGACGTCTACG	536
Db	392	CTTCGGAGAGATCTTCGCCCGACGACGAGATGCTGTGTGACGACGCCCATTC	439

```

RESULT 14
US-09-796-766-13
; Sequence 13, Application US/09796766
; Patent No. 6660850
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Lightner, Jonathan
; APPLICANT: Rafalecki, Antoni
; TITLE OF INVENTION: BRITTLE-1 HOMOLOGS
; FILE REFERENCE: BB157 US CIP
; CURRENT APPLICATION NUMBER: US/09/796,766
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/668884
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: PCT/US99/06583
; PRIOR FILING DATE: 1999-03-22
; PRIOR APPLICATION NUMBER: 60/079420
; PRIOR FILING DATE: 1998-03-26
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 13
; LENGTH: 1089
; TYPE: DNA
; ORGANISM: Glycine max
US-09-796-766-13

```

Query Match	1.9%	Score	42.2	DB	4	Length	1089
Best Local Similarity	70.9%	Pred. No.	0.25				
Matches	56	Conservative	0	Mismatches	23	Indels	0
						Gaps	0

Qy	2148	CACACCCTTAACTAATGCTCCCAAACTAATGAGTTAAATAAATACAATAACTCGTAAAAAA	2207
Dd	1011	CAGAGCTATATTATTTTTACAAATCAATTGTGTTTCATTAATAAAAAAAAAAAAAAAAAAAA	1070
Qy	2208	AAAAAAAAAAAAAAAAAAAAA	2226
Dd	1071	AAAAAAAAAAAAAAAAAAAAA	1089

RESULT 15
 US-09-252-991A-14038/C
 ; Sequence 14038; Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196, 136
 ; CURRENT APPLICATION NUMBER: US/09/252, 991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074, 788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094, 190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ. ID NOS: 33142
 ; SEQ ID NO 14038

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 04:26:28 ; Search time 742.548 Seconds
(without alignments)
11175.829 Million cell updates/sec

Title: US-09-622-964-2

Perfect score: 2229

Sequence: 1 cagsgagctccaccagccta.....aaaaaaaaaaaaaaaaaaaaa 2229

Scoring table: IDENTITY_NTC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

15: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1247.2	56.0	1263	10	US-09-746-783-3 Sequence 3, Appl1
2	587.8	26.4	1717	15	US-10-264-237-259 Sequence 259, App
3	459.2	20.6	2500	15	US-10-108-260A-1299 Sequence 1299, Ap
4	350.4	15.7	1326	13	US-10-071-766-10 Sequence 10, Appl
5	215.4	9.7	1350	9	US-09-768-826-16 Sequence 16, Appl
6	213	9.6	620	10	US-09-814-353-19504 Sequence 19504, A
7	189.2	8.5	1292	14	US-10-198-846-11070 Sequence 11070, A
8	175	7.9	1198	9	US-09-880-107-2174 Sequence 2174, Ap
9	161.4	7.2	578	15	US-10-027-633-134530 Sequence 134530,
10	149.4	6.7	748	15	US-10-027-633-134529 Sequence 134529,
11	135.2	6.1	539	14	US-10-029-386-7583 Sequence 7583, Ap
12	134.2	6.0	235	14	US-10-029-386-21283 Sequence 21283, A
13	116	5.2	853	14	US-10-198-846-11346 Sequence 11346, A
14	104	4.7	462	14	US-10-198-846-2561 Sequence 2561, Ap
15	103.6	4.6	3186778	15	US-10-027-632-174961 Sequence 174961,

16	102	4.6	615	14	US-10-198-846-7843	Sequence 7843, Ap
17	98	4.4	507	10	US-09-764-872-195	Sequence 195, App
18	66.2	3.0	442	10	US-09-814-353-14343	Sequence 14343, A
19	66	3.0	748	15	US-10-027-632-134529	Sequence 134529,
20	65.4	2.9	260	10	US-09-814-353-1601	Sequence 1601, Ap
21	65.4	2.9	260	10	US-09-814-353-7959	Sequence 7959, Ap
22	65.4	2.9	463	10	US-09-814-353-13563	Sequence 13563, A
23	64	2.8	751	14	US-10-198-846-1862	Sequence 1862, A
24	63.2	2.8	298	10	US-09-814-353-806	Sequence 806, App
25	63.2	2.8	298	10	US-09-814-353-7178	Sequence 7178, Ap
26	60	2.7	60	10	US-09-908-975-16358	Sequence 16358, A
27	54.4	2.4	65	10	US-09-908-975-23941	Sequence 23941, A
28	49.2	2.2	1263	10	US-09-746-783-3	Sequence 3, Appl1
29	49.2	2.2	1326	13	US-10-071-766-10	Sequence 10, Appl
30	48.2	2.2	8607	14	US-10-204-708-72	Sequence 72, Appl
31	48	2.2	16918	14	US-10-311-455-1589	Sequence 1589, Ap
32	47.8	2.1	2201	12	US-10-424-599-133253	Sequence 133253,
33	47.8	2.1	6668	14	US-10-311-455-1670	Sequence 1670, Ap
34	47.6	2.1	466	12	US-10-424-599-10863	Sequence 10863, A
35	47.4	2.1	12601	14	US-10-311-455-2179	Sequence 2179, Ap
36	47.2	2.1	891	14	US-10-156-761-6551	Sequence 6951, Ap
37	47.2	2.1	902568	14	US-10-156-761-1	Sequence 1, Appl
38	47	2.1	2076	10	US-09-986-480-116	Sequence 116, Appl
39	46.8	2.1	5349	14	US-10-311-455-313	Sequence 313, Appl
40	46.8	2.1	3673778	14	US-10-312-841-1	Sequence 1, Appl1
41	46.8	2.1	3673778	14	US-10-312-841-2	Sequence 25, Appl
42	46.6	2.1	846	9	US-09-349-015-25	Sequence 21, Appl
43	46.6	2.1	846	14	US-10-219-664-21	Sequence 7, Appl1
44	46.6	2.1	8195	9	US-09-795-693-7	
45	46.6	2.1	8195	14	US-10-156-239-7	

ALIGNMENTS

RESULT 1

US-09-746-783-3

Sequence 3, Application US/09746783

Publication No. US20030044935A1

GENERAL INFORMATION:

APPLICANT: Jacobe, Kenneth

McCoy, John M.

Ravallie, Edward R.

Racle, Lisa A.

Treacy, Maurice

Spaulding, Vikki

Agostino, Michael J.

Howes, Steven H.

Fechtel, Kim

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES

ENCODING THEM

NUMBER OF SEQUENCES: 231

CORRESPONDENCE ADDRESS:

ADDRESSER: Genetics Institute, Inc.

STREET: 87 Cambridgepark Drive

CITY: Cambridge

STATE: MA

COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/746,783

FILING DATE: 21-Dec-2000

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Milasincic, Debra J.

REGISTRATION NUMBER: 46,931

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

```

; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1263 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Query Match      56.0%; Score 1247.2; DB 10; Length 1263;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1249; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 978 GAGCAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGGAGACCACTGATTTGTC 1037
DB 1 GAGCAGCTCATCAACCCCTTTGGAGAGATGATGATTTTGGAGACCACTGATTTGTC 60

QY 1038 GACAGAAATTTGCAGGTGTCCTGTTGAGTGAGTGCACCGAGACCTTGCCTGCG 1097
DB 61 GACAGAAATTTGCAGGTGTCCTGTTGAGTGAGTGCACCGAGACCTTGCCTGCG 120

QY 1098 ATGAGCCGCGACATGTACTGGAATTAAGCCGAGCCACACAGCCCTTACACAGCTGCTTCC 1157
DB 121 ATGAGCCGCGACATGTACTGGAATTAAGCCGAGCCACAGCCCTTACACAGCTGCTTCC 180

QY 1158 GCCCAGTTCGCTGAGCCTCTCTTTATGAGGCTCCACCTTCAACATCAGCTGGAACAAAG 1217
DB 181 GCCCAGTTCGCTGAGCCTCTCTTTATGAGGCTCCACCTTCAACATCAGCTGGAACAAAG 240

QY 1218 GAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGAGATGCTCAAGCTGAGCATATTGGC 1277
DB 241 GAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGAGATGCTCAAGCTGAGCATATTGGC 300

QY 1278 CGCTTCTAGAGCCTGCAAGTCCCATGATCAACATCTTCCAGAGGCAAACTGAAGACCAA 1337
DB 301 CGCTTCTAGAGCCTGCAAGTCCCATGATCAACATCTTCCAGAGGCAAACTGAAGACCAA 360

QY 1338 CTACTGTGAGCCCAAGAGGAAATCCCTTCTCCAGAGAGGAGCCGCCCCAAACCAACAGGCA 1397
DB 361 CTACTGTGAGCCCAAGAGGAAATCCCTTCTCCAGAGAGGAGCCGCCCCAAACCAACAGGCA 420

QY 1398 GCCAACAAGAACGTTAAGGAGGCGCAGAGAGACAAACAGGCTGGAAGCTTAAAGCTGTGAG 1457
DB 421 GCCAACAAGAACGTTAAGGAGGCGCAGAGAGACAAACAGGCTGGAAGCTGTGAG 480

QY 1458 GCCTTCAAGTCTGGCCCACTGTATCAGAGGCGAGGCTTACTACAGTGCCTCCACAGAGCCGC 1517
DB 481 GCCTTCAAGTCTGGCCCACTGTATCAGAGGCGAGGCTTACTACAGTGCCTCCACAGAGCCGC 540

QY 1518 CTCAGCCCACTCCCATGTTCTTCCCTAGAAACATCAGGCGCGTCAAAAGCTTCAAGT 1577
DB 541 CTCAGCCCACTCCCATGTTCTTCCCTAGAAACATCAGGCGCGTCAAAAGCTTCAAGT 600

QY 1578 GTCAACAGGCAATGACACCAAAAGCAAAAGCTTAAAGACTGTGAGTTCTGGGCGCAAGAAA 1637
DB 601 GTCAACAGGCAATGACACCAAAAGCAAAAGCTTAAAGACTGTGAGTTCTGGGCGCAAGAAA 660

QY 1638 AGTTTGAATTTGCTCTCAGAGAGCGATGGGCGCTTATGAGAGACCCCAAGAAATTTCA 1697
DB 661 AGTTTGAATTTGCTCTCAGAGAGCGATGGGCGCTTATGAGAGACCCCAAGAAATTTCA 720

QY 1698 GTGAGAGAGAAAACGTGAGATTAACTGACGATATGACAGAGATCCCGAAATTCAC 1757
DB 721 GTGAGAGAGAAAACGTGAGATTAACTGACGATATGACAGAGATCCCGAAATTCAC 780

QY 1758 CTGAAGAAGACTTTGGAACATACCAACCAATACATACATCACTCACTCAAGATCAGT 1817
DB 781 CTGAAGAAGACTTTGGAACATACCAACCAATACATACATCACTCACTCAAGATCAGT 840

QY 1818 GATCCTTAATTGGGCTTTGGAAGAACAGGAGTGAAGCACTTCTTAACCTGCTTCAATAG 1877
```

```

DB 841 GATCCTTAATTGGGCTTTGGAAGAACAGGAGTGAAGCAATTCCTTAACCTGCTTCAATAG 900
QY 1878 GAATGCTTGCAGCAGCAGGAGCTCTACCTGTGTGTACACCGAGAGACACTGATCCAGTCA 1937
DB 901 GAATGCTTGCAGCAGCAGGAGCTCTACCTGTGTGTACACCGAGAGACACTGATCCAGTCA 960
QY 1938 CAGCCATACAGCTGTCCACACTGAAGAACGCTGTCTTAACAACGCTGAATCAATAGTTTA 1997
DB 961 CAGCCATACAGCTGTCCACACTGAAGAACATGTCTTAACAACGCTGAATCAATAGTTTA 1020
QY 1998 GCTTAATAGTAAATATCCAGACTCTCAGCTTAAATGCTTTATTCATTAATAACT 2057
DB 1021 GCTTAATAGTAAATATCCAGACTCTCAGCTTAAATGCTTTATTCATTAATAACT 1080
QY 2058 GTGAAGCTAGACTGAACCATTTGAACATTTAACTCAGACTCTGATTCAGAGTCGGGA 2117
DB 1081 GTGAAGCTAGACTGAACCATTTGAACATTTAACTCAGACTCTGATTCAGAGTCGGGA 1140
QY 2118 ACCCTTAAGTTCATCTGAATCCAGACAGCCACACCTTAAGTATCTGCCCAACTAATGA 2177
DB 1141 ACCCTTAAGTTCATCTGAATCCAGACAGCCACACCTTAAGTATCTGCCCAACTAATGA 1200
QY 2178 GTTAAATAATCAAAATACTCGTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 2229
DB 1201 GTTAAATAATCAAAATACTCGTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAA 1252

RESULT 2
US-10-264-237-259
; Sequence 259, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Biase et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; PRIOR FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 259
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1072)..(1072)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1596)..(1596)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1634)..(1634)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1657)..(1657)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1702)..(1702)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1712)..(1712)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-259
```

	Query Match Similarity	26.4%;	Score 587.8;	DB 15;	Length 1717;	
	Beet Local Similarity	98.8%;	Pred. No. 1.2e-157;			
	Matches	600;	Conservative	2;	Mismatches	4;
					Indels	1;
					Gaps	1
OY	1242	GAGACGAGAGAGTCTCACCCTGGCATCATTTGGCCCGTTTCCTAG;CCTGCAGTCCCAT	1301			
Db	1	GAGGACGAGAGAGATGTCTCAGCGTGGCATTTGGCCCGTTTCCTAG;CCTGCAGTCCCAT	60			
OY	1302	GATCACCATCTTCCCAAGGGCAAATCTCAGGCCAAACTAAGGCCAAACTAAGTGCCCAAGGGAATCC	1367			
Db	61	GATCACCATCTTCCCAAGGGCAAATCTCAGGCCAAACTAAGTGCCCAAGGGAATCC	120			
OY	1362	CTTCTCCACGAGGGGCTGCCAAAACCAAGGGAGCCAACAAGAGCTTAGGGGCCAG	1422			
Db	121	CTTCTCCACGAGGGGCTGCCAAAACCAAGGGAGCCAACAAGAGCTTAGGGGCCAG	179			
OY	1422	GAAGAACAACAGGCTTGAAAGCTTAAGGCTGTGAGAGCCTTCAAGTCTGGCCCACTGTAT	1483			
Db	180	GAAGAACAACAGGCTTGAAAGCTTAAGGCTGTGAGAGCCTTCAAGTCTGGCCCACTGTAT	239			
OY	1482	CAGAGGCCAGGCTACTACAGTGCCCCACAGAGGCCCTCAGCCCACTCCATGTTCTTC	1541			
Db	240	CAGAGGCCAGGCTACTACAGTGCCCCACAGAGGCCCTCAGCCCACTCCATGTTCTTC	299			
OY	1542	CCCCTGAAACCTACAGCCGCCCTCAAAAGTTCACTAGTTCACAGGCATAGACACCAAGAC	1601			
Db	300	CCCCTGAAACCTACAGCCGCCCTCAAAAGTTCACTAGTTCACAGGCATAGACACCAAGAC	359			
OY	1602	AAAACTTAAGAAGTGTGAGTTCTGTGGGGCCAAAGAAAGTTTGAATTGCTCAGAGAGC	1664			
Db	360	AAAACTTAAGAAGTGTGAGTTCTGTGGGGCCAAAGAAAGTTTGAATTGCTCAGAGAGC	419			
OY	1662	GATGGGGCCTTGATGAGAGACCCAGAAAGTATCTCAAGTAGAGAGAAAACTGTGAGATT	1721			
Db	420	GATGGGGCCTTGATGAGAGACCCAGAAAGTATCTCAAGTAGAGAGAAAACTGTGAGATT	479			
OY	1722	AACCTGAGGGATATGTCACAGATTCCTCCGAAAATCACTTCAAGAACCTTTGGAACAATCA	1781			
Db	480	AACCTGAGGGATATGTCACAGATTCCTCCGAAAATCACTTCAAGAACCTTTGGAACAATCA	539			
OY	1782	CCAACTCAATACATCACTACACTCACTCAAAAGATCACTAGATTCCTTATTTGGGCTTTGGAAC	1841			
Db	540	CCAACTCAATACATCACTACACTCACTCAAAAGATCACTAGATTCCTTATTTGGGCTTTGGAAC	599			
OY	1842	AGGAGTG 1848				
Db	600	AGGCTGTG 606				
	RESULT 3					
	US-10-108-260A-1299					
	; Sequence 1299, Application US//10108260A					
	; Publication No. US20040005560A1					
	; GENERAL INFORMATION:					
	; APPLICANT: HELIX RESEARCH INSTITUTE					
	; TITLE OF INVENTION: NO. US20040005560A1 full length cDNA					
	; FILE REFERENCE: H1-A0106					
	; CURRENT APPLICATION NUMBER: US/10/108,260A					
	; CURRENT FILING DATE: 2002-03-27					
	; NUMBER OF SEQ ID NOS: 5458					
	; SOFTWARE: PatentIn Ver. 2.1					
	; SEQ ID NO 1299					
	; LENGTH: 2500					
	; TYPE: DNA					
	; ORGANISM: Homo sapiens					
	US-10-108-260A-1299					
	Query Match	20.6%;	Score 459.2;	DB 15;	Length 2500;	
	Beet Local Similarity	63.5%;	Pred. No. 1.3e-120;			
	Matches	701;	Conservative	0;	Mismatches	403;
					Indels	0;
					Gaps	0;
OY	105	ATGACCATCATCTTAACAACAAGCCAAGTGGCTATGCGCCGTTAGGCTCTCTTCCCGCCTG	164			

Db	229	ATGACTGTCATTACTCCAGTAAGATGACAAATGCAACTTTTTTTGGATTTCATAGTTA	288
Oy	165	CTGCTGTGCTGCGCGGAGCATCTTACAGCTGCTAATGCGAGTTCTTAATCTTCCG	224
Db	289	CTCCCTCAAGATGGAGGAGCAGCATCTACAACTACTGTACAGGGAATTTATGTTTTGCT	348
Oy	225	CTCTGCTACTATCATCATCCGCTTTATTTATATAGGCTGGCCCTCACGGAAAGAACAAAGCTG	284
Db	349	GTTCTTTATACAGCAATTAAGTTGGTGTACAGATTGTTACTTACAGGAGTCCAAAAGCT	408
Oy	285	ATGTTTGGAAATGACTCTGTATTTGCGACAGCTACATCCAGCTCATCCCAATTTCTTTC	344
Db	409	TACTTTGAAAATTTATCAATTTTACTGTGACAGATATGCTGAAACAATTTCCAGTAACCTTT	468
Oy	345	GTCGTGGGCTTCTACGTGACCGCTGGTGTGACCCGCTGGTGGAAACAGTACGAAACCTG	404
Db	469	GTGCTTGGGTTTATATGTTACTCTGGTAGTGAACCGATGTTGGAAACCAATTTGTGAATTTG	528
Oy	405	CCGTGGCCCGACCGCCTCATATGAGCTGTGTCCGGCTTCTGTGAAGCGAAGACGACGAA	464
Db	529	CCCTGGCCGACAGAGCTTAATGTTCTCTCACTCTGACAGTGTTCACGGAAGGACGACGAC	588
Oy	465	GGCCGGCTGCTGGGCGGACGCTATCCGCTACAGCCAACTGGGCAAGTGTCTCATCTCG	524
Db	589	GGGCGCCTGCTTTAAGAAAGACCGCTGATGCGCTACGTCAATCTCACCCTCTCATCTTT	648
Oy	525	CGACGCTGACACCGCAGTCTACAGCGCTTCCCGACGCGCCGACACCTGGTGGACGA	584
Db	649	CGCTCCGATGACATCGTGTGTACAAAAGATTTCCCAATATGACACAGTGGTGAACGA	708
Oy	585	GGCTTTATGACTCCGGCAGAAACAGACAGTTGGAGAACTGAGCTTACACACAACTG	644
Db	709	GGTTTATGACAAACAGATGAAAGAAATTAATTCACCACTCAAGTCTCTCATCTGAAA	768
Oy	645	TTCTGGGGGCGCTGGGAGTGTGGTTGGCCAACTGTCAATGAAGGCGGTGGAGTGGGA	704
Db	769	TATTGGGTTCCATTCACTGTGTGTGGAAATTTTGCACTTAAGCCCGGAATGAAGGTGA	828
Oy	705	ATCCGGGACCCCTATCTGCTCCAGAGCGCTGTCAGACGATGAACACTTGTGCTACTG	764
Db	829	ATCAGAGACAGTGTGATCTGCAATCATATGATGACTGAATGAATGCATACCGCTCTGG	888
Oy	765	TGTGGACACTGTATGCTTACGCTACGACTGCAATTAATTAATCCACTGTGTATACACAGTGTG	824
Db	889	TGCAGCTCTTAATCCGTTATGACTGGGTTGGAAATCCGCTGTATACACCAGGTTGTC	948
Oy	825	ACTGTGGCGGTATACAGCTTCTTCTCTGACTTGTCTAGTTGGGCGCGAGTTTCTGAACCA	884
Db	949	ACTGTGTGTCTATACCTTCTTTTGGCGTGAATTTGAACGCGCAATTTTGTGATCCC	1008
Oy	885	GCCAAAGCCTTACCCTGGCCATGAGCTGGACCTGTTGTCGCGCTTTCACGTTCCCTGAG	944
Db	1009	ACCAAAAGCTTACGAGGAGGACATCTTGATTTTACATTTCCCATCTTCAACCCTCTACAA	1068
Oy	945	TTCTTTCTTATATGTTGCTGTGCTGAAGGTTGCGACAGACGCTCATCAACCCCTTTGGAGAG	1004
Db	1069	TTCTTTCTTATATGAGATGGCTTAAGGTAGCAAGACGCTTATCAACCCCTTTGGAGAA	1128
Oy	1005	GATATATGATATTTTGAACCAACTGTGATTTGTCAACAGAAATTTTGGAGAGTGTCCCTGTG	1066
Db	1129	GATATATGATATTTTGAACCTTAACTGTGTGATTTACAGAAATTTTGCAGGTCTCTTTTAA	1188
Oy	1065	GCTGTGATGATGATGACACAGGACCTGTGCTGTGATGAGCCGGAACATGTATCTGAAATTAAG	1124
Db	1189	GCTGTGACGAAATGTGACATGAGCTTACCCAAAGTGAAGAAAGACATTTTACTGGGACAT	1248
Oy	1125	CCCGAGCCACAGCCGCCCTTACACAGCTGTTCGCGCCAGTTCCGTGACGCTCTTATAG	1184
Db	1249	TCTGCTGCTGCGCCACCATACACTTTGGCAGCTGTGACTACTGCAATACCTCATTTCTG	1308
Oy	1185	GGCTTCACTTTCAACATCATGACCTG	1208
Db	1309	GGGTCAACAGTCCAGATGGGGCTG	1332

```
RESULT 4
US-10-071-766-10/c
; Sequence 10, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huel-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071.766
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match      15.7%; Score 350.4; DB 13; Length 1326;
Best Local Similarity 99.7%; Pred. No. 1.6e-89;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1850 AGCATTTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGTCTCTCACTGTGT 1909
    |||||
DB 1326 AGCATTTCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGCCAGGTCTCTCACTGTGT 1267

QY 1910 GTACACGACGACGACATGATCCAGTACAGCCATACAGCGTCCACATGAAAGACGTG 1969
    |||||
DB 1266 GTACACGACGACGACATGATCCAGTACAGCCATACAGCGTCCACATGAAAGACATG 1207

QY 1970 TCTTCAACAAGCTGGAATCAATGCTTAAATAGATAAAATCCACAGTACTTTCAG 2029
    |||||
DB 1206 TCTTCAACAAGCTGGAATCAATGCTTAAATAGATAAAATCCACAGTACTTTCAG 1147

QY 2030 CCTTAAATGCTTTATATCAATTAATAAAAGTGAAGCTAGACCTAACATGGAACATTT 2089
    |||||
DB 1146 CCTTAAATGCTTTATATCAATTAATAAAAGTGAAGCTAGACCTAACATGGAACATTT 1087

QY 2090 AACTCAGACTCTGGATTCAGAGTCGGGAAACCTTAGTTCATCTGAATCCAGACAGCCA 2149
    |||||
DB 1086 AACTCAGACTCTGGATTCAGAGTCGGGAAACCTTAGTTCATCTGAATCCAGACAGCCA 1027

QY 2150 CACCTTAGTATCTGCCCAAACTAATGATGTTAATAAATCAAAATCTGCTT 2201
    |||||
DB 1026 CACCTTAGTATCTGCCCAAACTAATGATGTTAATAAATCAAAATCTGCTT 975

RESULT 5
US-09-768-826-16
; Sequence 16, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PFS12PI
; CURRENT APPLICATION NUMBER: US/09/768.826
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
```

```
NAME/KEY: SITE
LOCATION: (1135)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1148)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1166)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1174)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1181)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1209)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1229)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1266)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1285)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1287)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1290)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1295)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1305)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1324)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1341)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,t,g, or c
US-09-768-826-16

Query Match      9.7%; Score 215.4; DB 9; Length 1350;
Best Local Similarity 61.9%; Pred. No. 9.6e-51;
Matches 420; Conservative 0; Mismatches 211; Indels 48; Gaps 3;

QY 603 GAACGACGACGCTGGAGAACTGAGCTACACACAACATGTTCTGGTGCCCTGGGTG 662
    |||||
DB 23 GAACGCTGGGGGTTTGAGAGCTGAATCCGACTTCAACAGTACTGGGTCCCTGGGTC 82

QY 663 TGGTTGGCAACCTGTCAATGAAGCGTGGCTTGAAGTGAATCCGGAGACCTATCCTG 722
    |||||
DB 83 TGGTTCAACAACCTGTGGCGGCCCAAGCCGAGAGGACGGGCAATACGTGACGATATCGCT 142

QY 723 CTCGAGAGCTGTGAACGAGATGAACACCTTGGCTACTGATGTGACACCTGTATGCC 782
    |||||
```

```
Db 143 CTCGTCTACTTTGGAGAGCTGAACAACTACCGACCAAGTGCACATGCTATTTCAC 202
Qy 783 TACAGCTGATAGTATCCCACTGATGTATACACAGGTGTGACTGNGCGGTGTACAG 842
Db 203 TATGACTGATCAGATCCCCCTCTGCTTACACCAAGTGTGACATAGCGCTACTCT 262
Qy 843 TTCTTCTGACTTGTCTAGTTGGCGGCACTTTCTGAAACCAAG-----885
Db 263 TTCTTTGCCCTCTCCCTGTTGGCCGCAATTTGTGAGCAGAGG-AGGGGCTGCCAA 322
Qy 886 -----CCAAAGCCTTACCTTGCCCATGAG---CTGAACTTC 917
Db 323 CCTCAGAGCTTCTGAAGCCAGGCGCAGAGCCCAAGCCTTGAGAGACCCGACATG 382
Qy 918 GTTGTGCGCGCTTTCACGCTTCTGACGTTCTTCTGATGTGGCTGCTGAGAGTGGCA 977
Db 383 TACGTGCTTTCACCACTGCTGCACTTCTTCTATGCTGCTGCTGCTGCAAGTGGCT 442
Qy 978 GAGCAGCTCATCAACCCCTTGGAGAGATGATGATTTTGAACCAACTGGAATGTC 1037
Db 443 GAAACATATCAACCACTTGTGTGAGATGATGAGACTTTGAGACAAATCAGCTCAT 502
Qy 1038 GACAGAAATTTGAGGTGTCCTGTTGCTGTGATGATGATGACCAAGACTGCTCGG 1097
Db 503 GACCGCAACTTGAGGATGCTCCTGCTATCCGTGAGAGAAATGTACAGAACTTCCCCC 562
Qy 1098 ATGAGACCGGACATGTATGGAATTAAGCCGACCAAGCCCTTACAC---AGTGTCT 1154
Db 563 GCTGAAGAGACCAAGTCTGGATGAGACCAAGCCGACCACTTACATGTGGACAG 622
Qy 1155 TCCGCCAGTTCCTGAGGCTCTTATAGGCTTCAACCTTCAACATCAGACTGAAACA 1214
Db 623 GCGCGGAGTCTGTGCGGCTTATTCCTGAGGCTTCAACCTTCAACCTGCGCATGAGCAG 682
Qy 1215 GAGAGATGAGATTCAGC 1233
Db 683 GACCTGAGACGAGCCTGC 701

RESULT 6
US-09-814-353-19504/c
; Sequence 19504, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814, 353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19504
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
```

```
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504
Query Match 9.6%; Score 213; DB 10; Length 620;
Best Local Similarity 66.4%; Pred. No. 2,9e-50;
Matches 306; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

Qy 103 CCATGACCACTTACAGAGCGAGTGAATGCGCGCTTACGCTCTTCCGCG 162
Db 464 CCATGACGGTTTCACTCTCAAGTGGAGGCGCGCTTCAAGGTTCTTGCC 405
Qy 163 TGTGCTGTGTCGCGCGGAGCATCTCAAGCTGTATATGCGAGTTCTTAATCTTC 222
Db 404 TGTCTTCCGCTGAGAGGGAAGATCTCAAGTCTCTTCAAGAAATCTCTCTTGG 345
Qy 223 TGTCTGTCTACATCATCCGCTTATTTATATATGCTGCGCTCAGCGAAGAACAGC 282
Db 344 GGGCTTGTATACGTGTGCTTAGCATCACTTACCGGCTGCTGTGACCCAGAGCAGAGT 285
Qy 283 TGAATTTGAGAACTGACTGTATTTGAGACGATCAATCCAGCTCATCCCATTTCT 342
Db 284 ACGTGTATCTAGGTGCGCCGCTGACCAACCGCTCAGACAGCTCATTTCTTGTCT 225
Qy 343 TGTGCTGCGGCTTCTACGTGAGCTGTGCTGACCGCTGTGTGAACAGTACAGAAC 402
Db 224 TGTATTTGGGTTTCTATATGATCTCTGTGTGAACCGCTGTGTGCTCCAGTACAGCA 165
Qy 403 TGTGCTGCGGACCGCTCATGAGCTGTGTGTGCGGCTTGTGAGAGCAAGACGAGC 462
Db 164 TCCCGGTGCAAGCAGCTGATGTGCGTATCTCGGCTAGCGGACCGCTGTGAGACGAG 105
Qy 463 AAGCGCGGCTGTGCGGCGGACGCTCATCCGCTTACCGCAACCTGCGCAAGCTCATCC 522
Db 104 GGGCGGCTGTGCGGCGGACCTCATCTGCTTACCGCAAGCTGTGTGTGTGTGTGTG 45
Qy 523 TGGCAGCGTCAAGACCGGAGTCTACAGCGCTTCCCGACG 563
Db 44 TGGCTCGGTCAAGACCGCGGCTGTATTAAGCGCTTCCCGACG 4

RESULT 7
US-10-198-846-11070
; Sequence 11070, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11070
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1287_1288, 1289, 1290, 1291, 1292
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

Query Match 8.5%; Score 189.2; DB 14; Length 1292;
Best Local Similarity 62.4%; Pred. No. 3.1e-43;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;
```

OY	349	GGGCTTTCTACTGAGACGCTGCTCGGACCCCGCTGGTGGAAACCAATGACAGGAACCTGCGCT	408
Db	318	TAGGGTTTTATGTACTCTGTGATGAAACCGAATGGTGAACCAATTGTGAATTTGCCCT	377
OY	409	GGCCCGACCGGCTCATAGAGCGCTGATCGAGGCTCTGTCGAAAGGCAAGACGACGAGCAAGCC	468
Db	378	GGCCGAGACAGGCTTAATGTTCTCATCTCTAGACAGTTCACCGGAACGACAGACACGGGG	437
OY	469	GGCTGCTGGGCGCAAGCTCATTCGGCTAAGCCAACTGGGCAACGTCGTCTATCTCTGGCA	528
Db	438	GCGTCTTTAGAAGAGACGCTGATGCGGTACGTCAATCTCACTCCGCTCATCTTTTCGCT	497
OY	529	GCGTCAGCACCGCAGCTCTACGACGCTTCCCCGACGCCAGCACCTGATGCAACGAGCT	588
Db	498	CGGTGAGCACTCTGTGTACAAAAGTTTCCACAAATGGAACACAGCTGTGGAAGCAGTT	557
OY	589	TTATGACTTCGGCAGAACCAAGCAGTGTGAGAACTGAGCCTTACCAACACATGTTCT	648
Db	558	TTATGACAAACAGATGAAGAAATTTTCAACACACTCAAGTCTCTCATCTGAAATATT	617
OY	649	GGGTGCCCCGGGTGTGTGTTCGCAACCTTCGATGAAGGCGTGGAGGTGCGAATCC	708
Db	618	GGTTTCCATTCACTGTGTGTGAAACTTTCGCACTTAAGCCGGAAATGAAGTAGAATCA	677
OY	709	GGGACCTTATCCTGTCTCCAGAGCCTGCTGACAGAGATGAAACCTTGCTACTCAGTGTG	768
Db	678	GAGACAGGTGTATCTGCAATCATTTGATGACTGAAATGAAATGCATACCGCTTTGGTGCA	737
OY	769	GACACCTGTATGCCCTACAGACTGGATTAAGTATCCCACTGGTGTATTAACAGAGTGG	822
Db	738	GCGCTTTATTCGTTATGACTGGGTGGGATTCGCTGTGTATTAACCCAGAGTGG	791

```

RESULT 8
US-09-880-107-2174/C
: Sequence 2174, Application US/09880107
: Patent No. US20020142981A1
: GENERAL INFORMATION:
: APPLICANT: Horne, Darci T.
: APPLICANT: Vockley, Joseph G.
: APPLICANT: Scherf, Uwe
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
: FILE REFERENCE: 44921-5028-WO
: CURRENT APPLICATION NUMBER: US/09/880,107
: CURRENT FILING DATE: 2001-06-14
: PRIOR APPLICATION NUMBER: US 60/211,379
: PRIOR FILING DATE: 2000-06-14
: PRIOR APPLICATION NUMBER: US 60/237,054
: PRIOR FILING DATE: 2000-10-02
: NUMBER OF SEQ ID NOS: 3950
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 2174
: LENGTH: 1198
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
US-09-880-107-2174
US-09-880-107-2174
US20020142981A1 L20941

```

	Query Match	7.9%;	Score 175;	DB 9;	Length 1198;
	Best Local Similarity	100.0%;	Pred. No. 3.6e-39;		
	Matches 175;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2027	CAGCCTTTAAAGCCTTTATTTCATATAAAACTGTGAACCTAGACTGAACCATTTGGAACA			20866
Db	1198	CAGCCTTTAAAGCCTTTATTTCATATAAAACTGTGAACCTAGACTGAACCATTTGGAACA			11399
QY	2087	TTTAACTCAGACTCTCGAGATTGAGACTGGGAAACCTTGGTCTATCTCAATCCAAAGACAG			21466
Db	1138	TTTAACTCAGACTCTCGAGATTGAGACTGGGAAACCTTGGTCTATCTCAATCCAAAGACAG			10799

Qy	2147	CCACACTTACTGCTGCCCAAACTAATGAGTTTAATAAATACCAAACTACTCGTT	2201
Db	1078	CCACACTTAGTATACCTGCCCAAAGTTTAATAAATACCAAACTACTCGTT	1024

```

RESULT 9
US-10-027-632-134530/c
; Sequence 134530, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 134530
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-134530

```

	Query Match	7.2%	Score 161.4	DB 15	Length 578	
	Best Local Similarity	99.4%	Pred. No. 1.8e-35			
	Matches	162	Conservative	0	Mismatches	1
					Indels	0
					Gaps	0
QY	582	GCAGGCTTTATGACCTCCGCGAGAAACACAAGCAGTGTGGAAACTGAGCTTACCAACAAC	641			
DB	183	GCAGGCTTTATGACTCCGCGAGAAACACAAGCAGTGTGGAAACTGAGCTTACCAACAAC	124			
QY	642	ATGTTCTTGAGTGCCCTCGGAGTGAGTTGCCAACCTGTCAATGAAGCGCTTGAGGT	701			
DB	123	ATGTTCTTGAGTGCCCTCGGAGTGAGTTGCCAACCTGTCAATGAAGCGCTTGAGGT	64			
QY	702	CGAATCCGGAGCCTTATCTCTGCTCCAGAGCCTGTGAACGAGA	744			
DB	63	CGAATCCGGAGCCTTATCTCTGCTCCAGAGCCTGTGAACGAGA	21			

```

RESULT 10
US-10-027-632-134529/c
Sequence 134529, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OR INVENTION: Identification and Mapping of Single Nucleotide
POLYMORPHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827,129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIORITY APPLICATION NUMBER: US 60/218,006
PRIORITY FILING DATE: 2000-07-12
PRIORITY APPLICATION NUMBER: US 60/198,676
PRIORITY FILING DATE: 2000-04-20
PRIORITY APPLICATION NUMBER: US 60/193,483
PRIORITY FILING DATE: 2000-03-29
PRIORITY APPLICATION NUMBER: US 60/185,218

```


;; PRIOR FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/167,363
;; PRIOR FILING DATE: 1999-11-23
;; PRIOR APPLICATION NUMBER: US 60/156,358
;; PRIOR FILING DATE: 1999-09-28
;; PRIOR APPLICATION NUMBER: US 60/146,002
;; PRIOR FILING DATE: 1999-08-09
;; NUMBER OF SEQ ID NOS: 325720
;; SOFTWARE: PatSeq for Windows Version 4.0
;; SEQ ID NO 134529
;; LENGTH: 748
;; TYPE: DNA
;; ORGANISM: Human
US-10-027-632-134529

Query Match 6.7%; Score 149.4; DB 15; Length 748;
Best Local Similarity 90.9%; Pred. No. 6.1e-32;
Matches 170; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 558 CCCAGGCGCCAGCAGCTGTGCAAGCGGCTTATGCTCCGGCAGACAGCAAGCAGTTG 617
DB 712 CCCAGCGCTTTTCACTCTCCACTCTGCAAGGCTTTATGACTCCGGCAGACAGCAG-1G 654
QY 618 GAGAACTGAGCCTACACACAAATGTTCTGGGCGCTGGTGTGTTGGCAACTG 677
DB 653 GAGAACTGAGCCTACACACAAATGTTCTGGGCGCTGGTGTGTTGGCAACTG 594
QY 678 TCATGAGCGCGTGGCTGTGAGTCCGAACTCCGGAGCCCTTATCTCTCTCCAGAGCTGCTG 737
DB 593 TCATGAGCGCGTGGCTGTGAGTCCGAACTCCGGAGCCCTTATCTCTCTCCAGAGCTGCTG 534
QY 738 AACGAG 744
DB 533 AACGTGA 527

RESULT 11

US-10-029-386-7583/C
; Sequence 7583, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 7583
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUE 9.00e-56
; OTHER INFORMATION: NT HIT: g115304094, EVALUE 1.00e-130
; OTHER INFORMATION: SWISSPROT HIT: O76090, EVALUE 2.00e-25
US-10-029-386-7583

Query Match 6.1%; Score 135.2; DB 14; Length 539;
Best Local Similarity 73.3%; Pred. No. 5.9e-28;
Matches 173; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
QY 351 GAGCTTCACTGACGCTGTGTCGACCCGCTGTGTGAACAGTAAGAACTGCGCTGG 410
DB 362 GCGTTTATGTGACGCTGTGTGTAACCGCTGTGTGAACAGTAAGTAATGATGCGCTG 303

QY 411 CCCAGCGCCTCATGAGCCTGTGTGTCGAGGCTTGTGCAAGGCAAGCAGCAGCCGG 470
DB 302 CCCAGCGCCTCATGAGCCTGTGTGTCGAGGCAAGCAGCAGCAGCAGCAGCCGG 243
QY 471 CTGCTCGGCGGCAAGCTCATCCGCTTACCGCAACTGCGGCAAGCTGTCTATCTGCGCAGC 530
DB 242 CTCTACCGGGGCACTCATGAGGCTTACCGCAGGAGCTCTCGGCGGTGCTCATCTGCGCTCC 183
QY 531 GTGAGCAGCGGCACTTACCAAGCGCTTCCCGAGGCCAGCAGCAGCTGTGTGCAAGAG 586
DB 182 GTGAGCAGCGGCTGTTCAGAGCGCTTCCCGACCATAGACCACTGTGTGAGGCTGG 127

RESULT 12

US-10-029-386-21283/C
; Sequence 21283, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 21283
; LENGTH: 235
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR19.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
; OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUE 4.00e-56
; OTHER INFORMATION: NT HIT: g18923136, EVALUE 1.00e-129
; OTHER INFORMATION: SWISSPROT HIT: O76090, EVALUE 1.00e-24
US-10-029-386-21283

Query Match 6.0%; Score 134.2; DB 14; Length 235;
Best Local Similarity 73.2%; Pred. No. 6.9e-28;
Matches 172; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 352 GCTTCACTGACGCTGTGTGTCGACCCGCTGTGTGGAACAGTAAGAACTGCGCTGGC 411
DB 235 GCTTATGTGAGCTGTGTGTGTAACCGCTGTGTGAGCCAGTAATGATATCCGCTGC 176
QY 412 CCGACCGCCTCATGAGCCTGTGTGTCGAGGCTTCTCGAAGGCAAGACAGAGCGCGGC 471
DB 175 CCGACCGCCTCATGAGCCTGTGTGTCGAGGCGCGGCAAGCGTGAACGAGCGAGCGCGGC 116
QY 472 TGCTGCGGCGCAGCTCATCCGCTTACCGCAACTGCGGCAAGCTGTCTCATCTCTGCGACGC 531
DB 115 TCAACCGGCGCAGCACTCATGCGCTTACGAGGCGCTCTCGGCGGCTCATCTCTGCGCTCCG 56
QY 532 TCAGCAGCGGCAAGCTTCAAGCGCTTCCCGAGGCCAGACAGCAGCTGTGTGCAAGCAG 586
DB 55 TCAGCAGCGGCGGTGTTCAGAGCGCTTCCCGACCATAGACCAAGTGTGTGAGGCTGG 1

RESULT 13

US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen

```
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11346
LENGTH: 853
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
LOCATION: 43, 50, 51, 54, 57, 853
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346
```

```
Query Match 5.2%; Score 116; DB 14; Length 853;
Best Local Similarity 68.2%; Pred. No. 2.6e-22;
Matches 161; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
```

```
QY 350 GGGCTTACGTACGCTGCTGCTGACCCGCTGCTGGAACCACTGACGAACCTGCTG 409
DB 272 GGGGTTTATGTACTCTGTGTAGAACCACTGCTGGAACCACTGCTGGAATTTGCGCTG 331
QY 410 GCCGACCGGCTCTATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 469
DB 332 GCCAGACAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 391
QY 470 GCTGCTGCGGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 529
DB 392 CCTGCTTGAAGAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 451
QY 530 CCGTACGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 585
DB 452 GGTGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 507
```

RESULT 14

```
US-10-198-846-2561/C
Sequence 2561, Application US/10198846
Publication No. US2003009974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2561
LENGTH: 462
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 390, 462
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-2561
```

```
Query Match 4.7%; Score 104; DB 14; Length 462;
```

```
Best Local Similarity 67.6%; Pred. No. 4.9e-19;
Matches 146; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 349 TGGGCTTACGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 408
DB 255 TAGGGTTTATGTACTCTGTGTAGAACCACTGCTGGAACCACTGCTGGAATTTGCGCT 196
QY 409 GGGCCGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 468
DB 195 GGCACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 136
QY 469 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 528
DB 135 GCGTGTGGAAGAAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 76
QY 529 GCGTACGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 564
DB 75 CCGTACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 40
```

RESULT 15

```
US-10-027-632-174961/C
Sequence 174961, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
POLYMOPLHISMS IN THE HUMAN GENOME
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 174961
LENGTH: 3186778
TYPE: DNA
ORGANISM: Human
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)...(3186778)
OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174961
```

```
Query Match 4.6%; Score 103.6; DB 15; Length 3186778;
Best Local Similarity 77.5%; Pred. No. 1.5e-16;
Matches 138; Conservative 0; Mismatches 39; Indels 1; Gaps 1;
```

```
QY 418 GCGTATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 477
DB 285842 GCGTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 285784
QY 478 GCGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 537
DB 285783 GGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285724
QY 538 CCGCAGTCTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 595
DB 285723 CCGTACGACCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 285666
```

Mon Mar 29 09:56:48 2004

us-09-622-964-2.rmpb

Page 9

Search completed: March 26, 2004, 11:54:34
Job time : 760.881 secs

This page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: March 25, 2004, 22:13:32 ; Search time 170.702 Seconds
(without alignments)
7896.639 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429

Sequence: 1 caggcgagcccccacagcccta.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued Patents NA: *
2: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfileseq1.seq: *

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	84.6	3.5	283	4 US-08-579-445-26	Sequence 26, Appl
2	84.6	3.5	5590	3 US-09-050-159-129	Sequence 129, App
3	84.6	3.5	9365	4 US-09-608-285A-8	Sequence 8, Appl
4	84.6	3.5	9365	4 US-09-350-836B-8	Sequence 8, Appl
5	84.6	3.5	9365	4 US-09-370-265-8	Sequence 8, Appl
6	84.6	3.5	9365	4 US-09-557-800C-8	Sequence 8, Appl
7	84.6	3.5	9365	4 US-09-370-625A-8	Sequence 8, Appl
8	84.6	3.5	14747	4 US-09-608-285A-42	Sequence 42, Appl
9	84.6	3.5	14747	4 US-09-557-800C-42	Sequence 42, Appl
10	84.6	3.5	15977	4 US-09-608-285A-59	Sequence 59, Appl
11	84.6	3.5	21784	4 US-09-820-002-3	Sequence 3, Appl
12	84.6	3.5	50000	4 US-09-146-053-3	Sequence 3, Appl
13	84.2	3.5	11531	1 US-08-068-945A-1	Sequence 1, Appl
14	84.2	3.5	11531	1 US-08-442-806-1	Sequence 1, Appl
15	84.2	3.5	11531	4 US-09-355-295B-1	Sequence 1, Appl
16	83.4	3.4	74962	4 US-09-685-853A-3	Sequence 3, Appl
17	83.4	3.4	786431	4 US-09-751-389-3	Sequence 3, Appl
18	83.2	3.4	36741	3 US-09-301-665-3	Sequence 3, Appl
19	83.2	3.4	41100	4 US-09-755-665-46	Sequence 46, Appl
20	83	3.4	319	4 US-09-621-976-13089	Sequence 13089, A
21	83	3.4	331	4 US-09-621-976-12648	Sequence 12648, A
22	83	3.4	1701	3 US-09-078-294-9	Sequence 9, Appl
23	83	3.4	2115	1 US-08-395-800A-7	Sequence 7, Appl
24	83	3.4	6769	1 US-08-480-784-20	Sequence 20, Appl
25	83	3.4	6769	1 US-08-483-553-20	Sequence 20, Appl
26	83	3.4	6769	1 US-08-487-002-20	Sequence 20, Appl
27	83	3.4	6769	1 US-08-483-554B-20	Sequence 20, Appl

28	83	3.4	6769	1 US-08-488-011B-20	Sequence 20, Appl
29	83	3.4	6769	3 US-08-850-727-20	Sequence 20, Appl
30	83	3.4	6769	5 PCT-US95-10202-20	Sequence 20, Appl
31	83	3.4	6769	5 PCT-US95-10203-20	Sequence 20, Appl
32	83	3.4	6769	5 PCT-US95-10220-20	Sequence 20, Appl
33	83	3.4	14636	3 US-09-173-914-6	Sequence 6, Appl
34	83	3.4	16891	4 US-09-486-147-1	Sequence 1, Appl
35	83	3.4	20674	4 US-09-641-638-651	Sequence 651, App
36	83	3.4	70000	4 US-09-851-896-3	Sequence 3, Appl
37	83	3.4	112132	4 US-09-741-150-3	Sequence 3, Appl
38	83	3.4	112132	4 US-10-160-187-3	Sequence 3, Appl
39	82.6	3.4	282	1 US-08-133-629-8	Sequence 8, Appl
40	82.6	3.4	1613	2 US-08-812-204-1	Sequence 1, Appl
41	82.6	3.4	8220	4 US-09-797-908-3	Sequence 3, Appl
42	82.6	3.4	11725	2 US-08-756-506-1	Sequence 1, Appl
43	82.6	3.4	11725	4 US-09-328-925-50	Sequence 50, Appl
44	82.6	3.4	18853	4 US-09-820-005-3	Sequence 3, Appl
45	82.6	3.4	38564	4 US-09-734-673-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-579-445-26
; Sequence 26, Application US/08579445
; Patent No. 6566053
; GENERAL INFORMATION:
; APPLICANT: Perucho, Manuel
; APPLICANT: Peinado, Miguel A.
; APPLICANT: Ionov, Yuri
; APPLICANT: Malkhosyan, Sergei
; TITLE OF INVENTION: Identification of Neoplasms by Detection
; TITLE OF INVENTION: Of Genetic Deletions
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESSES:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Releasee #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,445
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/152,484
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kirkpatrick, Anita M.
; REGISTRATION NUMBER: 32,617
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 283 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-579-445-26
Query Match 3.5%; Score 84.6; DB 4; Length 283;

```
Best Local Similarity 90.9%; Pred. No. 4,7e-12;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGGCGGTGGTGGCGCACCTGTAATCCAGCTACTCGGAGGCTGAGGC 1138
    |||||
DB 129 AAAAATTAGCGCGGCGGTGGTGGCGCGCTGTAAATCCAGCTACTCGGAGGCTGAGGC 188
    |||||

QY 1139 AGGAGATGCGCTTGAACCCGCGGAGCGGAGGCTTGTGGTG 1177
    |||||
DB 189 AGGAGATGCGCTTGAACCCGCGGAGCGGAGGCTTGTGGTG 227
    |||||

RESULT 2
US-09-050-159-129/c
; Sequence 129, Application US/09050159A
; Patent No. 6197505
; GENERAL INFORMATION:
; APPLICANT: No. 6197505berg, Leif T
; APPLICANT: Andersson, Maria K
; APPLICANT: Linstrom, Per H
; TITLE OF INVENTION: METHODS FOR ASSESSING CARDIOVASCULAR STATUS AND
; FILE REFERENCE: 1248/1D042
; CURRENT APPLICATION NUMBER: US/09/050,159A
; EARLIER FILING DATE: 1998-03-27
; EARLIER APPLICATION NUMBER: 60/042,930
; EARLIER FILING DATE: 1987-04-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 129
; LENGTH: 5590
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Angiotensin I converting enzyme, 5' region
US-09-050-159-129

Query Match
Best Local Similarity 90.9%; Score 84.6; DB 3; Length 5590;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGGCGGTGGTGGCGCACCTGTAATCCAGCTACTCGGAGGCTGAGGC 1138
    |||||
DB 1926 AAAAATTAGCGCGGCGGTGGTGGCGCGCTGTAAATCCAGCTACTCGGAGGCTGAGGC 1867
    |||||

QY 1139 AGGAGATGCGCTTGAACCCGCGGAGCGGAGGCTTGTGGTG 1177
    |||||
DB 1866 AGGAGATGCGCTTGAACCCGCGGAGCGGAGGCTTGTGGTG 1828
    |||||

RESULT 3
US-09-608-285A-8/c
; Sequence 8, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
; FILE REFERENCE: 28110/36570
; CURRENT APPLICATION NUMBER: US/09/608,285A
; EARLIER FILING DATE: 2000-06-30
; EARLIER APPLICATION NUMBER: 09/583,231
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
```

```
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (3409)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9214)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9303)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
; NAME/KEY: misc.feature
; LOCATION: (9311)
; OTHER INFORMATION: n = adenine or guanine or cytosine or thymidine
US-09-608-285A-8

Query Match
Best Local Similarity 84.8%; Score 84.6; DB 4; Length 9365;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGGCGGTGGTGGCGCACCTGTAATCCAGCTACTCGGAGGCTGAGGC 1138
    |||||
DB 7598 AAAAATTAGCGCGGCGGTGGTGGCGCGCTGTAAATCCAGCTACTCGGAGGCTGAGGC 7539
    |||||

QY 1139 AGGAGATGCGCTTGAACCCGCGGAGCGGAGGCTTGTGGTG 1177
    |||||
DB 7538 AGGAGATGCGCTTGAACCCGCGGAGCGGAGGCTTGTGGTG 7500
    |||||

RESULT 4
US-09-350-836B-8/c
; Sequence 8, Application US/09350836B
; Patent No. 6387645
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 9365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: exon
; LOCATION: (1)..(288)
```

```

? NAME/KEY: exon
? LOCATION: (1281)..(1580)
? NAME/KEY: exon
? LOCATION: (1820)..(1855)
? NAME/KEY: exon
? LOCATION: (2467)..(2555)
? NAME/KEY: exon
? LOCATION: (12863)..(2942)
? NAME/KEY: exon
? LOCATION: (3889)..(3950)
? NAME/KEY: exon
? LOCATION: (4894)..(4995)
? NAME/KEY: exon
? LOCATION: (5847)..(5987)
? NAME/KEY: exon
? LOCATION: (6966)..(7138)
? NAME/KEY: exon
? LOCATION: (8556)..(9365)
? NAME/KEY: misc feature
? LOCATION: (3405)
? OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
? NAME/KEY: misc feature
? LOCATION: (9214)
? OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
? NAME/KEY: misc feature
? LOCATION: (9303)
? OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
? NAME/KEY: misc feature
? LOCATION: (9311)
? OTHER INFORMATION: n= adenine or guanine or cytosine or thymine
? IS-09-350-836B-8

```

Query Match	3.5%	Score 84.6	DB 4	Length 9365
Best Local Similarity	84.8%	Pred. No. 3e-11		
Matches	84	Conservative	8	Mismatches 7; Indels 0; Gaps 0;
QY	1079	AGAAAGCAGCAGGGGTGTGGCGACACCTGTATCCAGCTCTGCGAGGCGTGAGGC	1138	
		: : : : : : : : : :		
Db	7598	AAAAATTAGCCRRGGGTGTGTGGCGRGVCCCTGTATCCAGCTACTTGGAGGCTGAGGC	7539	
QY	1139	AGGAGATCGCTGAACCCGGGAGCGGAGGTTGGTG	1177	
		: : : : : : : : :		
Db	7538	AGRAGATCGCTTGAAACCCAGAGAGGTGAGGTGACGTG	7500	

```

: RESULT 5
: US-09-370-265-8/c
: Sequence 8, Application US/09370265
: Patent No. 6447771
: GENERAL INFORMATION
: APPLICANT: Mulero, John
: APPLICANT: Mulero, Julio
: TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
: TITLE OF INVENTION: POLYPEPTIDES
: FILE REFERENCE: 2811/35908
: CURRENT APPLICATION NUMBER: US/09/370,265
: EARLIER FILING DATE: 1999-08-09
: EARLIER APPLICATION NUMBER: PCT/US99/16180
: EARLIER FILING DATE: 1999-07-16
: EARLIER APPLICATION NUMBER: 09/350,836
: EARLIER FILING DATE: 1999-07-09
: EARLIER APPLICATION NUMBER: 09/273,447
: EARLIER FILING DATE: 1999-03-19
: EARLIER APPLICATION NUMBER: 09/244,444
: EARLIER FILING DATE: 1999-02-04
: EARLIER APPLICATION NUMBER: 09/122,449
: EARLIER FILING DATE: 1998-07-24
: EARLIER APPLICATION NUMBER: 09/118,205
: EARLIER FILING DATE: 1998-07-16
: NUMBER OF SEQ ID NOS: 37
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 8
: LENGTH: 9365

```

```

TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (72)..(287)
FEATURE:
NAME/KEY: CDS
LOCATION: (1280)..(1579)
FEATURE:
NAME/KEY: CDS
LOCATION: (1819)..(1854)
FEATURE:
NAME/KEY: CDS
LOCATION: (2466)..(2555)
FEATURE:
NAME/KEY: CDS
LOCATION: (2863)..(2940)
FEATURE:
NAME/KEY: CDS
LOCATION: (3867)..(3952)
FEATURE:
NAME/KEY: CDS
LOCATION: (4896)..(4994)
FEATURE:
NAME/KEY: CDS
LOCATION: (5846)..(5986)
FEATURE:
NAME/KEY: CDS
LOCATION: (6965)..(7138)
FEATURE:
NAME/KEY: CDS
LOCATION: (8556)..(8639)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3409)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9214)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9303)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
FEATURE:
NAME/KEY: misc_feature
LOCATION: (9311)
OTHER INFORMATION: n = adenine or guanine or cytosine or thymine
US-09-370-265-8

```

```

Query Match      3.5% ; Score 84.6 ; DB 4 ; Length 9365 ;
Near Local Similarity 84.8% ; Prod. No. 3e-11 ;
Matches 84 ; Conservative 8 ; Mismatches 7 ; Indels 0 ; Gaps 0 ;

Oy    1079 AGAAGCAGCCAGGCGTGTGGCGCACCTGTATATCCACTACTCGGAGGCTGAGGC   1138
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db     7598 AAAATTATGACCRGGCGTGTTGGCCRRYGCCTGTATATCCCAAGCTACTYGGAGGCTGAGGC   7539

Oy    1139 AGAGATTCGCTTGAAACCCGGAGAGCCGAGAGTGTGTGTG   1177
      || : | | | | | | | | | | | | | | | | | | | | | | | |
Db     7538 AGRAGATTCCTTGTAACCAAGAGAGYRAGAAGTYTCAATG   7500

RESULT 6
US-09-557-800C-8/c
; Sequence 8, Application US/09557800C
; Patent No. 6476211
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides

```

```

: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 8
: LENGTH: 9365
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: exon
: LOCATION: (1)..(288)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (1281)..(1580)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (1820)..(1855)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (2467)..(2555)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (2863)..(2942)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (3889)..(3950)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (4894)..(4995)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (5847)..(5987)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (6966)..(7138)
: FEATURE:
: NAME/KEY: exon
: LOCATION: (8556)..(9365)
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (3409)
: OTHER INFORMATION: n = a or g or t or c
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (9214)
: OTHER INFORMATION: n = a or g or t or c
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (9303)
: OTHER INFORMATION: n = a or g or t or c
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (9311)
: OTHER INFORMATION: n = a or g or t or c
: OTHER INFORMATION: n = a or g or t or c
US-09-370-625A-8

Query Match      3.5%; Score 84.6; DB 4; Length 9365;
Best Local Similarity 84.8%; Pred. No. 3e-11; 7; Indels 0; Gaps 0;
Matches 84; Conservative 8; Mismatches

QY      1079 AGAAGAGCCAGGCGGTGTGGGCAACCTGTATCCGAGCTACTCGGAGGCTGAGGC 1138
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      7598 AAAATTATGCGCRGGCGGTGGTGGCGCGTGTATCCGAGCTACTYGGAGGCTGAGGC 7539
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY      1139 AGAGAAATCGCTTGAACCCGGGAGGCGGAGGTTGTGCTG 1177
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db      7538 AGRAGATTCCTTGAACCCGAGGAGGAGGTGACAGTG 7500
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 8
US-09-608-285A-42/c
: Sequence 42, Application US/09608285A
: Patent No. 6335013
: GENERAL INFORMATION:
: APPLICANT: Ford, John

```



```
/ APPLICANT: Mulero, Julio
/ APPLICANT: Yeung, George
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
/ FILE OF INVENTION: POLYPEPTIDES
/ FILE REFERENCE: 28110/36570
/ CURRENT APPLICATION NUMBER: US/09/608,285A
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: 09/583,231
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 09/557,800
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/481,238
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: 09/370,265
/ PRIOR FILING DATE: 1999-08-09
/ PRIOR APPLICATION NUMBER: PCT/US99/16180
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: 09/350,836
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 09/273,447
/ PRIOR FILING DATE: 1999-03-19
/ PRIOR APPLICATION NUMBER: 09/244,444
/ PRIOR FILING DATE: 1999-02-04
/ PRIOR APPLICATION NUMBER: 09/122,449
/ PRIOR FILING DATE: 1998-07-24
/ PRIOR APPLICATION NUMBER: 09/118,205
/ PRIOR FILING DATE: 1998-07-16
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 42
/ LENGTH: 14747
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (13641)
/ OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-608-285A-42

Query Match      3.5%; Score 84.6; DB 4; Length 14747;
Best Local Similarity 84.8%; Pred. No. 3.8e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGAGCGGTGGCGCACACCTGTAAATCCGAGTACTGGAGGCTGAGGC 1138
DB 10958 AAAAATTAGCGCGCGCTGTGGCGCGCGCTGTAAATCCGAGTACTGGAGGCTGAGGC 10899
QY 1139 AGGAGATCGCTTGAACCCGAGGCGGAGGTTGTGCTG 1177
DB 10898 AGRAGAATCGCTTGAACCCGAGGCGGAGGTTGTGCTG 10860

RESULT 9
US-09-557-800C-42/c
/ Sequence 42, Application US/09557800C
/ Patent No. 6476211
/ GENERAL INFORMATION:
/ APPLICANT: Ford, John
/ APPLICANT: Mulero, Julio
/ APPLICANT: Yeung, George
/ TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
/ FILE REFERENCE: 28110/36457
/ CURRENT APPLICATION NUMBER: US/09/557,800C
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/481,238
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: 09/370,265
/ PRIOR FILING DATE: 1999-08-09
/ PRIOR APPLICATION NUMBER: PCT/US99/16180
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: 09/350836
/ PRIOR FILING DATE: 1999-07-09
```

```
/ PRIOR APPLICATION NUMBER: 09/273447
/ PRIOR FILING DATE: 1999-03-19
/ PRIOR APPLICATION NUMBER: 09/122449
/ PRIOR FILING DATE: 1998-07-24
/ PRIOR APPLICATION NUMBER: 09/244444
/ PRIOR FILING DATE: 1999-02-04
/ PRIOR APPLICATION NUMBER: 09/118,205
/ PRIOR FILING DATE: 1998-07-16
/ NUMBER OF SEQ ID NOS: 56
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 42
/ LENGTH: 14747
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (13641)
/ OTHER INFORMATION: n = adenosine or guanine or cytosine or thymidine
US-09-557-800C-42

Query Match      3.5%; Score 84.6; DB 4; Length 14747;
Best Local Similarity 84.8%; Pred. No. 3.8e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGAGCGGTGGCGCACACCTGTAAATCCGAGTACTGGAGGCTGAGGC 1138
DB 10958 AAAAATTAGCGCGCGCTGTGGCGCGCGCTGTAAATCCGAGTACTGGAGGCTGAGGC 10899
QY 1139 AGGAGATCGCTTGAACCCGAGGCGGAGGTTGTGCTG 1177
DB 10898 AGRAGAATCGCTTGAACCCGAGGCGGAGGTTGTGCTG 10860

RESULT 10
US-09-608-285A-59/c
/ Sequence 59, Application US/09608285A
/ Patent No. 6335013
/ GENERAL INFORMATION:
/ APPLICANT: Ford, John
/ APPLICANT: Mulero, Julio
/ APPLICANT: Yeung, George
/ TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE
/ FILE REFERENCE: 28110/36570
/ CURRENT APPLICATION NUMBER: US/09/608,285A
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: 09/583,231
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 09/557,800
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/481,238
/ PRIOR FILING DATE: 2000-01-11
/ PRIOR APPLICATION NUMBER: 09/370,265
/ PRIOR FILING DATE: 1999-08-09
/ PRIOR APPLICATION NUMBER: PCT/US99/16180
/ PRIOR FILING DATE: 1999-07-16
/ PRIOR APPLICATION NUMBER: 09/350,836
/ PRIOR FILING DATE: 1999-07-09
/ PRIOR APPLICATION NUMBER: 09/273,447
/ PRIOR FILING DATE: 1999-03-19
/ PRIOR APPLICATION NUMBER: 09/244,444
/ PRIOR FILING DATE: 1999-02-04
/ PRIOR APPLICATION NUMBER: 09/122,449
/ PRIOR FILING DATE: 1998-07-24
/ PRIOR APPLICATION NUMBER: 09/118,205
/ PRIOR FILING DATE: 1998-07-16
/ NUMBER OF SEQ ID NOS: 60
/ SOFTWARE: Patentln Ver. 2.0
/ SEQ ID NO 59
/ LENGTH: 15977
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
```

OTHER INFORMATION: CD39-14/L66 Gene Sequence
NAME/KEY: CDS
LOCATION: (245)..(461)
NAME/KEY: CDS
LOCATION: (1454)..(1533)
NAME/KEY: CDS
LOCATION: (2734)..(2877)
NAME/KEY: CDS
LOCATION: (4364)..(4439)
NAME/KEY: CDS
LOCATION: (4679)..(4714)
NAME/KEY: CDS
LOCATION: (5326)..(5414)
NAME/KEY: CDS
LOCATION: (5723)..(5802)
NAME/KEY: CDS
LOCATION: (6751)..(6812)
NAME/KEY: CDS
LOCATION: (7758)..(7859)
NAME/KEY: CDS
LOCATION: (8712)..(8852)
NAME/KEY: CDS
LOCATION: (9831)..(9887)
NAME/KEY: CDS
LOCATION: (11613)..(11728)
NAME/KEY: CDS
LOCATION: (13146)..(13691)
NAME/KEY: CDS
LOCATION: (15702)..(15839)
NAME/KEY: misc_feature
LOCATION: (14871)
OTHER INFORMATION: n = a or c or g or t
US-09-608-285A-59

Query Match 3.5%; Score 84.6; DB 4; Length 15977;
Best Local Similarity 84.8%; Pred. No. 4e-11;
Matches 84; Conservative 8; Mismatches 7; Indels 0; Gaps 0;
QY 1079 AGAAGACGCGCGTGTGGCGCACCTGTAATCCGACTACCGGAGGCTGAGGC 1138
DB 12188 AAAAATTAGCCRCGGCGTGGCGCGCTGTAAATCCGACTACGAGGCTGAGGC 12129
QY 1139 AGGAGATCGCTTGAACCCGAGCGGAGGTTGTGATG 1177
DB 12128 AGRAGAATCGCTTGAACCCGAGGAGTGRAGGTGCAGTG 12090

RESULT 11
US-09-820-002-3/C
Sequence 3, Application US/09820002
Patent No. 6482630
GENERAL INFORMATION:
APPLICANT: Gan, Weiniu
APPLICANT: Ye, Jane
APPLICANT: DiFrancesco, Valentina
APPLICANT: Beasley, Ellen
TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
FILE REFERENCE: CL001194
CURRENT APPLICATION NUMBER: US/09/820,002
CURRENT FILING DATE: 2001-03-29
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PASCSEQ for Windows Version 4.0
SEQ ID NO 3
LENGTH: 21784
TYPE: DNA
ORGANISM: HUMAN
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(21784)
OTHER INFORMATION: n = A,T,C or G
US-09-820-002-3

Query Match 3.5%; Score 84.6; DB 4; Length 21784;
Best Local Similarity 83.5%; Pred. No. 4.7e-11;
Matches 96; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 1079 AGAAGACGCGCGTGTGGCGCACCTGTAATCCGACTACTCGGAGGCTGAGGC 1138
DB 20871 AAAAATCAGCTGGGCGTGGTGGGCGACCGCTGTAATCCGACTACTCGGAGGCTGAGGC 20812
QY 1139 AGGAGATCGCTTGAACCCGAGCGGAGGTTGTGTCAGAGCCTCATCA 1193
DB 20811 AGGAGATCGCTTGAACCCGAGGCGGAGGTTGTAGTGAAGCTAGATACGCCA 20757

RESULT 12
US-09-146-053-3/C
Sequence 3, Application US/09146053A
Patent No. 639349
GENERAL INFORMATION:
APPLICANT: Ryan, James W.
APPLICANT: Sprinkle, Terry Joe Curtis
APPLICANT: Venema, Richard C
TITLE OF INVENTION: Human Amino-peptidase P Gene
FILE REFERENCE: MCG103
CURRENT APPLICATION NUMBER: US/09/146,053A
CURRENT FILING DATE: 1998-09-02
EARLIER APPLICATION NUMBER: 60/057,854
EARLIER FILING DATE: 1997-09-02
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 50000
TYPE: DNA
ORGANISM: Homo sapiens
US-09-146-053-3

Query Match 3.5%; Score 84.6; DB 4; Length 50000;
Best Local Similarity 90.9%; Pred. No. 7.2e-11;
Matches 90; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
QY 1079 AGAAGACGCGCGTGTGGCGCACCTGTAATCCGACTACTCGGAGGCTGAGGC 1138
DB 8689 AAAAATTAGCCCAAGGTGTGGCGCACCGCTGTAATCCGACTACTCGGAGGCTGAGGC 8630
QY 1139 AGGAGATCGCTTGAACCCGAGCGGAGGTTGTGATG 1177
DB 8629 AGGAGATCGCTTGAACCCGAGGAGTGGAGTTGCAGTG 8591

RESULT 13
US-08-068-945A-1
Sequence 1, Application US/08068945A
Patent No. 5616483
GENERAL INFORMATION:
APPLICANT: Bjursell, Gunnar
APPLICANT: Carlsson, Peter
APPLICANT: Enerback, Sven
APPLICANT: Hansson, Lemnat
APPLICANT: Lidberg, Ulf
APPLICANT: Nilsson, Jeanette
APPLICANT: Tornell, Jan
TITLE OF INVENTION: New DNA Sequences
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSER: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/068,945A
FILING DATE: 27-MAY-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201809-2
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201826-6
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9202088-2
FILING DATE: 03-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9300902-5
FILING DATE: 19-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)819-8783
TELEFAX: (212)354-8113
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11531 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Mammary gland
FEATURE:
NAME/KEY: CDS
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
OTHER INFORMATION: /BC_number=3.1.1.1
FEATURE:
NAME/KEY: /product= "Bile Salt-Scimulated Lipase"
LOCATION: 1..1640
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1611..1617
FEATURE:
NAME/KEY: exon
LOCATION: 1641..1727
FEATURE:
NAME/KEY: exon
LOCATION: 4071..4221
FEATURE:
NAME/KEY: exon
LOCATION: 4307..4429
FEATURE:
NAME/KEY: exon
LOCATION: 4707..4904
FEATURE:
NAME/KEY: exon
LOCATION: 6193..6323
FEATURE:
NAME/KEY: exon
LOCATION: 6501..6608
FEATURE:
NAME/KEY: exon
LOCATION: 6501..6608
FEATURE:
NAME/KEY: exon

LOCATION: 6751..6868
FEATURE:
NAME/KEY: exon
LOCATION: 8335..8521
FEATURE:
NAME/KEY: exon
LOCATION: 8719..8922
FEATURE:
NAME/KEY: exon
LOCATION: 10124..10321
FEATURE:
NAME/KEY: exon
LOCATION: 10650..11490
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 11491..11531
US-08-068-945A-1
Query Match 3.5%; Score 84.2; DB 1; Length 11531;
Best local Similarity 91.8%; Pred. No. 4.2e-11;
Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1081 AAGCAGCCAGCGGTGTGGCCGACACTGTATCCAGCTACTCGGAGGCTGAGGCAG 1140
|||
Db 5361 AATTAGCCAGCGGTGTGGCCGCTTGTATCCAGCTACTCGGAGGCTGAGGCAG 5420
QY 1141 GAGATCGCTTGAACCCGGAGCGGAGGTTGTGCTG 1177
|||
Db 5421 GAGATCGCTTGAACCTCAGAGGCGGAGGTTGCGGTG 5457
|||
RESULT 14
US-08-442-806-1
Sequence 1, Application US/08442806
Patent No. 5716817
GENERAL INFORMATION:
APPLICANT: Bjursell, Gunnar
APPLICANT: Carlsson, Peter
APPLICANT: Enerback, Sven
APPLICANT: Hansson, Lennart
APPLICANT: Lidberg, Ulf
APPLICANT: Nilsson, Jeanette
APPLICANT: Tornell, Jan
TITLE OF INVENTION: Genomic DNA Sequences
TITLE OF INVENTION: Encoding Human BSSL/CEL
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSER: White & Case
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: United States
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/442,806
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/068,945
FILING DATE: 27-MAY-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201809-2
FILING DATE: 11-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9201826-6
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA:

APPLICATION NUMBER: SE 9202088-2
FILING DATE: 03-JUL-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: SE 9300902-5
FILING DATE: 19-MAR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Steiner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-052
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 11531 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Mammary gland
FEATURE:
NAME/KEY: CDS
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
FEATURE:
NAME/KEY: mat_peptide
LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
OTHER INFORMATION: /BC_number= 3.1.1.1
OTHER INFORMATION: /product= "Bile Salt-Stimulated Lipase"
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..1640
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1611..1617
FEATURE:
NAME/KEY: exon
LOCATION: 1641..1727
FEATURE:
NAME/KEY: exon
LOCATION: 4071..4221
FEATURE:
NAME/KEY: exon
LOCATION: 4307..4429
FEATURE:
NAME/KEY: exon
LOCATION: 4707..4904
FEATURE:
NAME/KEY: exon
LOCATION: 6193..6323
FEATURE:
NAME/KEY: exon
LOCATION: 6501..6608
FEATURE:
NAME/KEY: exon
LOCATION: 6751..6868
FEATURE:
NAME/KEY: exon
LOCATION: 8335..8521
FEATURE:
NAME/KEY: exon
LOCATION: 8719..8922
FEATURE:
NAME/KEY: exon
LOCATION: 10124..10321
FEATURE:
NAME/KEY: exon
LOCATION: 10650..11490

FEATURE:
NAME/KEY: 3'UTR
LOCATION: 11491..11531
US-08-442-806-1
Query Match 3.5%; Score 84.2; DB 1; Length 11531;
Best Local Similarity 91.8%; Pred. No. 4.2e-11;
Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1081 AAAGACCGCAGCGCTGTGGCCACACCTGTATCCGAGCTATCTGGAGGCTGAGGCAG 1140
DB 5361 AAATGACGACGCGGTGGCGCTTGCCTGTATCCGAGCTATCTGGAGGCTGAGGCAG 5420
QY 1141 GAGATCGCTTGAACCGGAGCGGAGGATTTGTGTTG 1177
DB 5421 GAGATCGCTTGAACCTCAGAGGCGGAGGTTGGTGTG 5457
RESULT 15
US-09-355-295B-1
Sequence 1, Application US/09355295B
Patent No. 6525241
GENERAL INFORMATION:
APPLICANT: Dairymple, M.
APPLICANT: Lundberg, L.
APPLICANT: Stromqvist, M.
TITLE OR INVENTION: Expression Methods
FILE REFERENCE: 1754 SEQUENCE LISTING v2a.txt
CURRENT APPLICATION NUMBER: US/09/355,295B
PCT/SE99/00648
CURRENT FILING DATE: 1999-07-18
PRIOR APPLICATION NUMBER: PCT/SE99/00648
PRIOR FILING DATE: 1999-04-21
NUMBER OF SEQ ID NOS: 5
SEQ ID NO 1
LENGTH: 11531
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genomic DNA
FEATURE:
OTHER INFORMATION: Mammary gland source
FEATURE:
NAME/KEY: CDS
LOCATION: join(1653..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11394)
NAME/KEY: mat_peptide
LOCATION: join(1722..1727, 4071..4221, 4307..4429, 4707
LOCATION: ..4904, 6193..6323, 6501..6608, 6751..6868, 8335
LOCATION: ..8521, 8719..8922, 10124..10321, 10650..11391)
OTHER INFORMATION: /BC_number= 3.1.1.1
OTHER INFORMATION: /product= "Bile Salt-Stimulated Lipase"
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..1640
FEATURE:
NAME/KEY: TATA_signal
LOCATION: 1611..1617
FEATURE:
NAME/KEY: exon
LOCATION: 1641..1727
FEATURE:
NAME/KEY: exon
LOCATION: 4071..4221
FEATURE:
NAME/KEY: exon
LOCATION: 4307..4429
FEATURE:
NAME/KEY: exon
LOCATION: 4707..4904
FEATURE:
NAME/KEY: exon

```

/ LOCATION: 6193..6323
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 6501..6608
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 6751..6868
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 8335..8521
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 8719..8922
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 10124..10321
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: 10650..11490
/ FEATURE:
/ NAME/KEY: 3'UTR
/ LOCATION: 11491..11531
US-09-355-295B-1

```

```

Query Match 3.5%; Score 84.2; DB 4; Length 11531;
Best Local Similarity 91.8%; Pred. No. 4.2e-11;
Matches 89; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1081 AAGCAGCCAGCGCTGTGCGCAGACCTGTAAATCCAGCTACTCCGAGGCTGAGGCAG 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5361 AATTAGCCAGCGCTGTGCGCCTGTGCTGTAAATCCAGCTACTCCGAGGCTGAGGCAG 5420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 1141 GAGATCGCTTGAAACCGGAGGCGAGGCTTGCTG 1177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 5421 GAGATCGCTTGAACTAGAGGCGAGGCTTGCGTG 5457
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: March 26, 2004, 05:11:11
 Job time : 173.702 sec

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 26, 2004, 04:26:28 ; Search time 809.174 Seconds
(without alignments)

11175.829 Million cell updates/sec

Title: US-09-622-964-4

Perfect score: 2429

Sequence: 1 caggagagcccccacagcctc.....aaaaaaaaaaaaaaaaaaaaa 2429

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: Published Applications NA.*
2: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10A_NEW_PUB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10C_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1244.2	51.2	1263	10	US-09-746-783-3	Sequence 3, Appl1
2	587.8	24.2	1717	15	US-10-264-237-259	Sequence 259, App
3	350.4	14.4	1326	13	US-10-071-766-10	Sequence 10, Appl
4	338.2	13.9	2500	15	US-10-108-260A-1299	Sequence 1299, App
5	213	8.8	620	10	US-09-814-353-19504	Sequence 19504, A
6	189.2	7.8	1292	14	US-10-198-846-11070	Sequence 11070, A
7	175	7.2	1198	9	US-09-880-107-2174	Sequence 2174, Ap
8	161.4	6.6	578	15	US-10-027-632-134530	Sequence 134530,
9	149.4	6.2	748	15	US-10-027-632-134529	Sequence 134529,
10	135.2	5.6	539	14	US-10-029-386-7583	Sequence 7583, Ap
11	134.2	5.5	235	14	US-10-029-386-21283	Sequence 21283, A
12	126.8	5.2	1350	9	US-09-768-826-16	Sequence 16, Appl
13	116	4.8	853	14	US-10-198-846-11346	Sequence 11346, A
14	104	4.3	462	14	US-10-198-846-2561	Sequence 2561, Ap
15	103.6	4.3	3186778	15	US-10-027-632-174961	Sequence 174961,

16	102	4.2	615	14	US-10-198-846-7843	Sequence 7843, Ap
17	98	4.0	507	10	US-09-764-872-195	Sequence 195, App
18	90	3.7	624	15	US-10-027-632-243509	Sequence 243509,
19	90	3.7	684973	9	US-09-263-959-1	Sequence 1, Appl1
20	89	3.7	31994	9	US-09-764-860-599	Sequence 599, App
21	89	3.7	31994	9	US-09-764-904-71	Sequence 71, Appl
22	89	3.7	31994	14	US-10-091-548-71	Sequence 599, App
23	89	3.7	31994	14	US-10-074-095-599	Sequence 599, App
24	89	3.7	31994	15	US-10-212-877-599	Sequence 11074, A
25	88.4	3.6	723	15	US-10-027-632-11024	Sequence 118248,
26	88.4	3.6	727	15	US-10-027-632-148248	Sequence 148249,
27	88.4	3.6	727	15	US-10-027-632-148249	Sequence 280638,
28	88.2	3.6	549	15	US-10-027-632-280638	Sequence 3851, Ap
29	88	3.6	30568	9	US-09-764-877-3851	Sequence 3851, Ap
30	88	3.6	30568	15	US-10-242-515-3851	Sequence 15, Appl
31	87.8	3.6	291	14	US-10-229-058B-15	Sequence 55069, A
32	87.8	3.6	625	15	US-10-027-632-55059	Sequence 55069, A
33	87.8	3.6	625	15	US-10-027-632-55060	Sequence 293995,
34	87.8	3.6	625	15	US-10-027-632-293995	Sequence 823, App
35	87.8	3.6	625	15	US-10-027-632-823	Sequence 823, App
36	87.8	3.6	895	15	US-10-027-632-823	Sequence 2049, Ap
37	87.8	3.6	895	15	US-10-027-632-823	Sequence 1695, Ap
38	87.8	3.6	38918	14	US-10-017-161-2049	Sequence 1, Appl1
39	87.8	3.6	38918	15	US-10-292-798-1695	Sequence 195571,
40	87.8	3.6	13393	13	US-10-161-510-1	Sequence 21091, A
41	87.4	3.6	468	15	US-10-027-632-195571	Sequence 99425, A
42	87.4	3.6	522	15	US-10-027-632-21091	Sequence 2903, Ap
43	87.4	3.6	616	15	US-10-027-632-99425	Sequence 2903, Ap
44	87.4	3.6	4963	9	US-09-764-877-2903	Sequence 2903, Ap
45	87.4	3.6	4963	15	US-10-242-515-2903	Sequence 2903, Ap

ALIGNMENTS

RESULT 1
US-09-746-783-3
; Sequence 3, Application US/09746783
; Publication No. US20030044935A1
; GENERAL INFORMATION:
; APPLICANT: Jacoby, Kenneth
; McCoy, John M.
; LaValle, Edward R.
; Racie, Lisa A.
; Treacy, Maurice
; Spaulding, Vikki
; Agostino, Michael J.
; Howes, Steven H.
; Fechtel, Kim
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; NUMBER OF SEQUENCES: 231
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent-In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/746,783
; FILING DATE: 21-Dec-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Milasincic, Debra J.
; REGISTRATION NUMBER: 46,931
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400

```

RESULT 2
US-10-264-237-259
: Sequence 259, Application US/10264237
: Publication No. US20040009451A1
GENERAL INFORMATION:
APPLICANT: Birst et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P411P1
CURRENT APPLICATION NUMBER: US/10/264,237
CURRENT FILING DATE: 2002-10-04
PRIOR APPLICATION NUMBER: PCT/US01/16450
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/205,515
NUMBER OF SEQ ID NOS: 2876
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 259
LENGTH: 1717
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1072)..(1072)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1596)..(1596)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1634)..(1634)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1657)..(1657)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1702)..(1702)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1712)..(1712)
OTHER INFORMATION: n equals a,t,g, or c
US-10-264-237-259

```


Query Match 24.2%; Score 587.8; DB 15; Length 1717;
Best Local Similarity 98.8%; Pred. No. 2.3e-151;
Matches 600; Conservative 2; Mismatches 4; Indels 1; Gaps 1;

QY 1445 GAGGACGAGAGAGATCTCAACGCTGGCATTCATTGGCCGCTTCTTACGCTGCACTGATCCAT 1504
DB 1 GAGGACGAGAGAGATCTCAACGCTGGCATTCATTGGCCGCTTCTTACGCTGCACTGATCCAT 60
QY 1505 GATTCACATCTCTCCAGGGGCAAACTCAAGGACCAAACTACTGTGGGCGCCAAAGGGGAATCC 1564
DB 61 GATTCACATCTCTCCAGGGGCAAACTCAAGGACCAAACTACTGTGGGCGCCAAAGGGGAATCC 120
QY 1565 CTCTCTCCAGAGGGGCTGCGCCAAACCAAGGCGAGCCAAACAGAACTTTAGGGGCGAG 1624
DB 121 CTCTCTCCAGAGGGGCTGCGCCAAACCAAGGCGAGCCAAACAGAACTTTAGGGGCGAG 179
QY 1625 GAAAGCAACAAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCACTGTAT 1684
DB 180 GAAAGCAACAAAGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGAGCCCACTGTAT 239
QY 1685 CAGAGGCGAGGCTACTCAAGTGGCCCAAGACGCGCCCTCAAGCCCACTCCCATGTTCTTC 1744
DB 240 CAGAGGCGAGGCTACTCAAGTGGCCCAAGACGCGCCCTCAAGCCCACTCCCATGTTCTTC 299
QY 1745 CCCCTAGAACATCAAGCGCGCTCAAGCTTCAAGTGTCAAGGCAATAGACACCAAGAC 1804
DB 300 CCCCTAGAACATCAAGCGCGCTCAAGCTTCAAGTGTCAAGGCAATAGACACCAAGAC 359
QY 1805 AAAAGCTTAAAGCTGTGAGTCTGCGGCGCCAAAGAAAGTTTGAATTCCTTCAGAGAC 1864
DB 360 AAAAGCTTAAAGCTGTGAGTCTGCGGCGCCAAAGAAAGTTTGAATTCCTTCAGAGAC 419
QY 1865 GATGGGCGCTTATGTGAGGACCCAGAAAGTATCTCAAGTGAAGGAGAAACTGTGAGATT 1924
DB 420 GATGGGCGCTTATGTGAGGACCCAGAAAGTATCTCAAGTGAAGGAGAAACTGTGAGATT 479
QY 1925 AACCTGACCGATATGCGAGAGATCCCGGAAATCACTTCAAGAACTTTGGAAACATCA 1984
DB 480 AACCTGACCGATATGCGAGAGATCCCGGAAATCACTTCAAGAACTTTGGAAACATCA 539
QY 1985 CCAACCAACATACACACTCACTCAAAAGTCAATGATCTTATTTGAGCCTTGGAAAAAC 2044
DB 540 CCAACCAACATACACACTCACTCAAAAGTCAATGATCTTATTTGAGCCTTGGAAAAAC 599
QY 2045 AAGGATG 2051
DB 600 AAGTCTG 606

RESULT 3
US-10-071-766-10/c
; Sequence 10, Application US/10071766
; Publication No. US20020192678A1
; GENERAL INFORMATION:
; APPLICANT: Huel-Mei Chen
; TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
; FILE REFERENCE: PA-0043 US
; CURRENT APPLICATION NUMBER: US/10/071,766
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match 14.4%; Score 350.4; DB 13; Length 1336;
Best Local Similarity 99.7%; Pred. No. 7e-86;
Matches 351; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2053 AGACATTCCTTAACCTGCTTCTTAATGGGATGCTTGGCAGCGAGGTCTCAGCTGTG 2112
DB 1326 AGACATTCCTTAACCTGCTTCTTAATGGGATGCTTGGCAGCGAGGTCTCAGCTGTG 1267
QY 2113 GTACACAGCAGAGACATGATCAAGTCAAGCCATACAGCTGTCCACTGAAAGACGT 2172
DB 1266 GTACACAGCAGAGACATGATCAAGTCAAGCCATACAGCTGTCCACTGAAAGACGT 1207
QY 2173 TCTTACAAAGCTGATGAATGATGTAGCTTATATGATTAATAATCCAGACTACTTCAG 2232
DB 1206 TCTTACAAAGCTGATGAATGATGTAGCTTATATGATTAATAATCCAGACTACTTCAG 1147
QY 2233 CCTTAATGCTTTTATTCATTAATAAATGTAAGAGTACAGTCAATGGAACATTT 2292
DB 1146 CCTTAATGCTTTTATTCATTAATAAATGTAAGAGTACAGTCAATGGAACATTT 1087
QY 2293 AACTCAGACTGTGATTCAGAGTGGGAAACCTTATGTTCTATCTGAATCCAGACGCA 2352
DB 1086 AACTCAGACTGTGATTCAGAGTGGGAAACCTTATGTTCTATCTGAATCCAGACGCA 1027
QY 2353 CACTTAGTATCTGCGCCAAACTAATGATTTAATTAATCAATCTGTT 2404
DB 1026 CACTTAGTATCTGCGCCAAACTAATGATTTAATTAATCAATCTGTT 975

RESULT 4
US-10-108-260A-1299
; Sequence 1299, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1299
; LENGTH: 2500
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1299

Query Match 13.9%; Score 338.2; DB 15; Length 2500;
Best Local Similarity 61.8%; Pred. No. 2.3e-82;
Matches 538; Conservative 0; Mismatches 333; Indels 0; Gaps 0;

QY 105 ATGACCATCACTTACCAAGCAAGGAGGCTTAAGCCGCTTAGGCTCTTCCCGCGCTG 164
DB 229 ATGACCTGTACTTACTCTCCAGTAAAGTAGCAATGCACTTTTGGATTTCAATAGGTTA 288
QY 165 CTGCTGTGCTGGCGGGGAGCATTTACAAAGCTGTATATGGGAGTTCTTAAATCTTCTG 224
DB 289 CTCTCAAGTGAAGGAGGAGCATTTCAAAACTAGTACAGGGAATTTATGTTTGGCT 348
QY 225 CTCTGCTACTACATATCCGCTTATTTATATAGCTGGCCCTCAAGGAAACACAGCTG 284
DB 349 GTTCTTTATACAGCAATTAATGTTGGTGTACAGATGTTACTTACAGGAGTCCAAAAAGT 408
QY 285 ATGTTGGAACCTAGCTGTATGCGAGCAGCTACATCAGCTCATGCCAATTCCTTC 344
DB 409 TACTTGAATAATTTATCAATTTTACTGTGACAGATATGCTGAACAAATTCAGTAACCTTT 468
QY 345 GTGCTGGCTTCTAGAGTACGCTGTGATCGTGAACCGCTGTGTGAACCAAGTACAGAACTG 404
DB 469 GTGCTGGCTTCTAGAGTACGCTGTGATCGTGAACCGATGTGTGAACCAAGTGTGAAATTTG 528
QY 405 CCGTGCCCGAGCCGCTCATAGAGCTGTGTGGGCTTGTGAAAGCAGAGCAGAGCA 464
DB 529 CCGTGCCCGAGAGGCTATATGTTCTCATCTTACAGAGTTCACAGGAACGACGAGCAGC 588
QY 465 GCGCGGCTGTGCGGCGAGCGCTCATCCGCTACGCGCAACCTGGGCAACGTGCTATCTG 524

```

Db      589 GGGGCCCTGCTTAGAAGACCGTGTATGCGTACATCTTCCACTCCCTGCTATCTTT 648
Qy      525 GCGACGTCAGACCGCAGTCTACAAAGCGCTTCCCGACGCCCGACCTGGTGCAAGCA 584
Db      649 CCGCTGCTGACGACTGCTGTGTACAAAGATTCTCCCAATATGACCAAGTGTGAAAGCA 708
Qy      585 GCGCTTATGACTCCGCGAAGACAAAGAGAGTTGAGAAATCTGAGCTTACCAAGACATG 644
Db      709 GGTATTATGACAAAGATGAAGAAATTTATTCACCACTTCAAGTCTCTCTATCTGAAA 768
Qy      645 TTCTGGGTGCGCGGTGTGTGTTGCTCAACTGTCAATGAAGCGGTGGCTTGAAGTCCA 704
Db      769 TATGAGTTTCAATTCACTGTGTTTGAATCTTGAACCTTAAAGCCCGAATGAAGTAGA 828
Qy      705 ATCCGGAGACCTATCTCTGCTCCAGACCTGCTGAACGAGATGAACACTTGGTACTAG 764
Db      829 ATCAGAGACAGTGTATGATCTGCAATCATATGATGACGAATGAATGATATACCGCTTGG 888
Qy      765 TGTGACACCTGTATGCTATGACGATGATGATATGATATGATATGATATGATATGATG 824
Db      889 TGCAGCCTTATTCGTTATGCTATGATGATGATGATGATGATGATGATGATGATG 948
Qy      825 ACTGTGCGGTGTACAGCTTCTCTGACTTGTCTAGTTGGCGGAGATTTCTGAACCA 884
Db      949 ACTTGTGCTGTATACCTTCTTCTTGTGCGTGTATGAGCGCAGATTTTGTGATCC 1008
Qy      885 GCCAAGGCTTACCTGCGCATGAGCTGAGCTGTTGCGGCTTACAGTCTCTGAG 944
Db      1009 ACCAAGGCTTACGAGGAGGAGTGTGATCTTACATTCATTCACCTCTTACAA 1068
Qy      945 TTCTTCTTATGTGTGCTGCTGAAGGTG 975
Db      1069 TTCTTCTTATGAGGATGCTTAAGTAG 1099

```

```

RESULT 5
US-09-814-353-19504/c
; Sequence 19504, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lee, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19504
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504

```

```

Query Match      8.8%; Score 213; DB 10; Length 620;
Best Local Similarity 66.4%; Pred. No. 4e-48; Indels 0; Gaps 0;
Matches 306; Conservative 0; Mismatches 155;
Qy      103 CCATGACATCACTTACACAAAGTGTATATGCTTATGAGCTTCTTCTCCGCC 162
Db      464 CCAAGACGTTTATATACATCTTCAAGGCGGAGCGCTTCCGAGGTTCTTGCC 405
Qy      163 TGTGCTGTGTGCGCGGAGCATATTAAGCTGTATATGAGCGGATTTATCTTCC 222
Db      404 TGTCTCCGCTGAGGAGGAGATCTACAGCTCTCTTCAAGAAATTCCTCTTTG 345
Qy      223 TGTCTGTATCACTATCATATCCGTTATTTATATAGCTGACCGCTTCAAGAAACACAGC 282
Db      344 GGGCTGTGACGCTGTGCTTATGATACCTACCGGCTGCTGACCCGAGAGAGGT 285
Qy      283 TGAATGTTGAAATATGACTCTGTATTTGAGACAGCTATCATCCAGCTATCCCATTTCT 342
Db      284 ACGTATATGCTCAGGTGCGCGGTACTGCAACCGCTCAGACAGCTCATATTCCTTCT 225
Qy      343 TGTGCTGTGCTTCTATGATGATGATGATGATGATGATGATGATGATGATGATG 402
Db      224 TTGTATGAGGTTTCTATGATATCTTGTGTGAACCGTGTGTCCAGTACACAAGA 165
Qy      403 TGGCGTGGCCGACCGGCTCATAGAGCTGTGTGCTGCTTCTGCAAGCAAGAGAGAG 462
Db      164 TCCCGCTGCGACGACCACTGATGATGATGATGATGATGATGATGATGATGATG 105
Qy      463 AAGCGCGGCTGTCTGCGGCGACGCTCATCCGCTACGCAACCTGAGGCAAGCTCATCC 522
Db      104 GGGCGCGGCTGTCTGCGGCGACCTCATCCGCTACGCAACCTGAGGCAAGCTGTGTGTG 45
Qy      523 TGGCGAGGCTGACGACCGGCTTACAGGCGTTTCCCGAG 563
Db      44 TGGCTGTGATGACACCGCGTGTGTTAAGCGCTTCCCGAC 4

```

```

RESULT 6
US-10-198-846-11070
; Sequence 11070, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11070
; LENGTH: 1292
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1287, 1288, 1289, 1290, 1291, 1292
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070

```

```

Query Match      7.8%; Score 189.2; DB 14; Length 1292;
Best Local Similarity 62.4%; Pred. No. 2.2e-41;
Matches 296; Conservative 0; Mismatches 178; Indels 0; Gaps 0;
Qy      349 TGGGCTTCTACGTCAGCTGTGTGACCCGCTGTGGAACAGTACGAGACCTCCGT 408

```

Db 318 TAGGGTTTATGTACTGTAGTGAACCGATGTGAAACCAAGTTGTGATTTGCCCT 317
Qy 409 GGGCCGACGGCCCTCATGAGCTGTGTGCGGGCTTTCGCGAAGGCAAGGACGAGGCC 468
Db 378 GGGCAGACAGGCTTAATGTTCTCTATCTTAGAGGTGTTCAAGAGCGAGACGAGGCG 437
Qy 469 GGGTGTGGGGCGGCGACGCTCATCGGTAGCGCAACCTGTGGCAAGTCTCATCTCGCA 528
Db 438 GCGCTGTTAGAGAGCGCTGATCGGTAGCTCAATCTCCCTGCTCATCTTTCCGCT 497
Qy 529 GCGTCAGCACCGGAGCTGTACAGCGCTTCCCGAGCGCGCCAGCACTGTGTGACAGCGCT 588
Db 498 CCGTGAGCACTGTGTGTACAAAGATTTCCCAATGTGACCGTGTGTAAGCAGGTT 557
Qy 589 TTATGACTCCGGGCAACCAAGCAGTGTGGAAAGTGAAGCCCTACGACCAATGTTCT 648
Db 558 TTATGCAACAGTGAAGAAATTTATCAACCACTCAAGTCTCTCATCTGAATATTT 617
Qy 649 GGGTGGCTGGGTGTGTGTCGCAACCTGTCAATGAAGCGGTGTGGAGTCAATCC 708
Db 618 GGGTTCATTCACTGTGTGTGAAATCTTGCAACTAAAGCCGGAATGAAGTGAATCA 677
Qy 709 GGGACCTTATCTGCTTCCAGAGCTGCTGAAAGATGAACACCTTGCTACTCACTGTG 768
Db 678 GAGACAGTGTGTGATCTGCAATCATGTGATGATGAATGAATCATACCGCTTGTGTGA 737
Qy 769 GACACCTGTATGCTGCTGAGATTAATGATCCCACTGTGTGTATACAGAGTGTG 822
Db 738 GCGCTTATGCTGTGTATGAGCTGGGTGTGGAATTCGCTGTATACACCGAGTAG 791

RESULT 7

US-09-880-107-2174/c
Sequence 2174, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
PRIOR FILING DATE: 2000-10-02
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2174
LENGTH: 1198
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L:0941
US-09-880-107-2174

Query Match 7.2%; Score 175; DB 9; Length 1198;
Best Local Similarity 100.0%; Pred. No. 1.7e-37;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2230 CAGCCTTAAATGCTTTTATTTCTAATAAACTGTGAAAGCTAGACTGAACATTGGAAAC 2289
Db 1198 CAGCCTTAAATGCTTTTATTTCTAATAAACTGTGAAAGCTAGACTGAACATTGGAAAC 1139
Qy 2290 TTTAATCAGACTCTGATTCAGAGTCGGGAAACCTTAGTCTATCTGATCCAAAGACG 2349
Db 1138 TTTAATCAGACTCTGATTCAGAGTCGGGAAACCTTAGTCTATCTGATCCAAAGACG 1079
Qy 2350 CCAGACCTTAGTACTGCCCAACTAATGATTTAATAATAACAATACTCGTT 2404
Db 1078 CCAGACCTTAGTACTGCCCAACTAATGATTTAATAATAACAATACTCGTT 1024

RESULT 8

US-10-027-632-134530/c
Sequence 134530, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134530
LENGTH: 578
TYPE: DNA
ORGANISM: Human
US-10-027-632-134530

Query Match 6.6%; Score 161.4; DB 15; Length 578;
Best Local Similarity 99.4%; Pred. No. 6.7e-34;
Matches 162; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 582 GCAAGCTTATGACTCCGGCAGAACAGACAGTGTGAGAAATGAGCTTACCAACAC 641
Db 183 GCAAGCTTATGACTCCGGCAGAACAGACAGTGTGAGAAATGAGCTTACCAACAC 124
Qy 642 ATGTTCTGGGTGCTCCGGGTGTGTTGGCAACCTGCAATGAAGCGGTGTGAGGT 701
Db 123 ATGTTCTGGGTGCTCCGGGTGTGTTGGCAACCTGCAATGAAGCGGTGTGAGGT 64
Qy 702 CGAATCCGGGACCTATCTGCTCCAGAGCTGCTGAACGAG 744
Db 63 CGAATCCGGGACCTATCTGCTCCAGAGCTGCTGAACGTGA 21

RESULT 9

US-10-027-632-134529/c
Sequence 134529, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358
 PRIOR FILING DATE: 1999-09-28
 PRIOR APPLICATION NUMBER: US 60/146,002
 PRIOR FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FASTSEQ for Windows Version 4.0
 SEQ ID NO: 134529
 LENGTH: 748
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-134529

Query Match 6.2%; Score 149.4; DB 15; Length 748;
 Best Local Similarity 90.9%; Pred. No. 1.6e-30;
 Matches 170; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

QY 558 CCCAGGCGCCAGACCTGCTGGAAGAGGCTTATGACTCCGCGAGAACACAGCACTTG 617
 712 CCCAGCGCTTTTCACTCCACTCTGCAAGCTTTATGACTCCGCGAGAACACAGCACTG 654
 QY 618 GAGAACTGAGCTTACCAACCAACATGTTCTGCGTCCCTGGGTGTTGCGCAACCTG 677
 DB 653 GAGAACTGAGCTTACCAACCAACATGTTCTGCGTCCCTGGGTGTTGCGCAACCTG 694
 QY 678 TCAATGAAAGCGTGTGCTTGAAGGTCGGAATCCGGAGCCCTATCCTGCTCCAGAGCTGCTG 737
 DB 593 TCAATGAAAGCGTGTGCTTGAAGGTCGGAATCCGGAGCCCTATCCTGCTCCAGAGCTGCTG 534
 QY 738 AACGAGA 744
 DB 533 AACGAGA 527

RESULT 10
 US-10-029-386-7583/c
 Sequence 7583; Application US/10029386
 Publication No. US20030194704A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Hanel, David K.
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
 TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 FILE REFERENCE: AEWICA-X-2
 CURRENT APPLICATION NUMBER: US/10/029.386
 CURRENT FILING DATE: 2001-12-20
 NUMBER OF SEQ ID NOS: 34286
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 7583
 LENGTH: 539
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE: MAP TO CHR19.3
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
 OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALU 9.00e-56
 OTHER INFORMATION: NT HIT: g15304094, EVALU 1.00e-130
 OTHER INFORMATION: SWISSPROT HIT: O76090, EVALU 2.00e-25
 US-10-029-386-7583

Query Match 5.6%; Score 135.2; DB 14; Length 539;
 Best Local Similarity 73.3%; Pred. No. 1.1e-26;
 Matches 173; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 351 GCTTCTACGCTGAGCTGTGACCGCTGTGTGAACCACTGCGAAGAACTGCGCTGG 410
 DB 362 GCTTCTACGCTGAGCTGTGACCGCTGTGTGAACCACTGCGAAGAACTGCGCTGG 303
 QY 411 CCCAGCGCTTATGAGCGCTGTGCTGCGGCTTCTGCAAGGCAAGCAAGCAAGCGCG 470

DB 302 CCCAGCGCTTATGAGCGCTGTGCTGCGGCTTCTGCAAGGCAAGCAAGCAAGCGCGCG 243
 QY 471 CTGCTGCGCGCGACGCTTATGCTGCTACCGCACTGCGGCAAGCTGCTATCTGCGCAGC 530
 DB 242 CTGCTGCGCGCGACGCTTATGCTGCTACCGCACTGCGGCTTCTGCGGCTGCTATCTGCGC 183
 QY 531 GTGAGCAACCGCAAGCTTACCAAGCGCTTCCGCGAGCCCGCAAGCACTGTGCAAGCAG 586
 DB 182 GTGAGCAACCGCGGCTTACCAAGCGCTTCCGCGAGCCCGCAAGCACTGTGCAAGCAG 127

RESULT 11
 US-10-029-386-21283/c
 Sequence 21283; Application US/10029386
 Publication No. US20030194704A1
 GENERAL INFORMATION:
 APPLICANT: Penn, Sharon G.
 APPLICANT: Hanel, David K.
 TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
 TITLE OF INVENTION: EXPRESSION ANALYSIS TWO
 FILE REFERENCE: AEWICA-X-2
 CURRENT APPLICATION NUMBER: US/10/029.386
 CURRENT FILING DATE: 2001-12-20
 NUMBER OF SEQ ID NOS: 34288
 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
 SEQ ID NO: 21283
 LENGTH: 235
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE: MAP TO CHR19.3
 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
 OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALU 4.00e-56
 OTHER INFORMATION: NT HIT: g18923136, EVALU 1.00e-129
 OTHER INFORMATION: SWISSPROT HIT: O76090, EVALU 1.00e-24
 US-10-029-386-21283

Query Match 5.5%; Score 134.2; DB 14; Length 235;
 Best Local Similarity 73.2%; Pred. No. 1.3e-26;
 Matches 172; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 352 GCTTCTACGCTGAGCTGTGCTGACCGCTGTGGAACAGTACGAACTGCTGCTGGC 411
 DB 235 GCTTCTATGAGCGCTGTGCTGGAACCGCTGTGGAACAGTACGAACTGCTGCTGGC 176
 QY 412 CGAGCGGCTTATGAGCGCTGTGCTGGAACCGCTGTGGAACAGTACGAACTGCTGCTGGC 471
 DB 175 CGAGCGGCTTATGAGCGCTGTGCTGGAACCGCTGTGGAACAGTACGAACTGCTGCTGGC 116
 QY 472 TGTGCGGCGGCAAGCTTATGCTGCTGCAAGCGGCAAGTGTCTATCTGCGCAGCG 531
 DB 115 TTAACGCGGCGCAAGCTTATGCTGCTGCAAGCGGCTTCTGCGGCTGCTATCTGCGCAGCG 56
 QY 532 TGAAGCAACCGAGCTTACCAAGCGCTTCCGCGAGCCCGCAAGCACTGTGCAAGCAG 586
 DB 55 TGAAGCAACCGAGCTTACCAAGCGCTTCCGCGAGCCCGCAAGCACTGTGCAAGCAG 1

RESULT 12
 US-09-768-826-16
 Sequence 16; Application US/09768826
 Patent No. US20020012966A1
 GENERAL INFORMATION:
 APPLICANT: Shi et al.
 TITLE OF INVENTION: 18 human secreted proteins
 FILE REFERENCE: PFS12P1
 CURRENT APPLICATION NUMBER: US/09/768.826
 CURRENT FILING DATE: 2001-01-25
 PRIOR APPLICATION NUMBER: PCT/US00/22350

```

1  PRIOR FILING DATE: 2000-08-15
2  PRIOR APPLICATION NUMBER: 60/148,759
3  PRIOR FILING DATE: 1999-08-16
4  NUMBER OF SEQ ID NOS: 61
5  SOFTWARE: PatentIn Ver. 2.0
6  SEQ ID NO 16
7  LENGTH: 1350
8  TYPE: DNA
9  ORGANISM: Homo sapiens
10 FEATURE:
11 NAME/KEY: SITE
12 LOCATION: (1135)
13 OTHER INFORMATION: n equals a,t,g, or c
14 NAME/KEY: SITE
15 LOCATION: (1148)
16 OTHER INFORMATION: n equals a,t,g, or c
17 NAME/KEY: SITE
18 LOCATION: (1166)
19 OTHER INFORMATION: n equals a,t,g, or c
20 NAME/KEY: SITE
21 LOCATION: (1174)
22 OTHER INFORMATION: n equals a,t,g, or c
23 NAME/KEY: SITE
24 LOCATION: (1181)
25 OTHER INFORMATION: n equals a,t,g, or c
26 NAME/KEY: SITE
27 LOCATION: (1209)
28 OTHER INFORMATION: n equals a,t,g, or c
29 NAME/KEY: SITE
30 LOCATION: (1229)
31 OTHER INFORMATION: n equals a,t,g, or c
32 NAME/KEY: SITE
33 LOCATION: (1266)
34 OTHER INFORMATION: n equals a,t,g, or c
35 NAME/KEY: SITE
36 LOCATION: (1285)
37 OTHER INFORMATION: n equals a,t,g, or c
38 NAME/KEY: SITE
39 LOCATION: (1287)
40 OTHER INFORMATION: n equals a,t,g, or c
41 NAME/KEY: SITE
42 LOCATION: (1290)
43 OTHER INFORMATION: n equals a,t,g, or c
44 NAME/KEY: SITE
45 LOCATION: (1295)
46 OTHER INFORMATION: n equals a,t,g, or c
47 NAME/KEY: SITE
48 LOCATION: (1305)
49 OTHER INFORMATION: n equals a,t,g, or c
50 NAME/KEY: SITE
51 LOCATION: (1324)
52 OTHER INFORMATION: n equals a,t,g, or c
53 NAME/KEY: SITE
54 LOCATION: (1339)
55 OTHER INFORMATION: n equals a,t,g, or c
56 NAME/KEY: SITE
57 LOCATION: (1341)
58 OTHER INFORMATION: n equals a,t,g, or c
59 NAME/KEY: SITE
60 LOCATION: (1343)
61 OTHER INFORMATION: n equals a,t,g, or c
62 NAME/KEY: SITE
63 LOCATION: (1345)
64 OTHER INFORMATION: n equals a,t,g, or c
65 NAME/KEY: SITE
66 LOCATION: (1348)
67 OTHER INFORMATION: n equals a,t,g, or c
68 US-09-768-826-16
69
70 Query Match 5.2% Score 126.8; DB 9; Length 1350;
71 Best Local Similarity 69.9%; Pred. No. 3,8e-24;
72 Matches 186; Conservative 0; Mismatches 77; Indels 3; Gaps 1

```

QY	1174	GGTGGCAGAGAGGCTCATCAACCCCTTTGGAGAGATGATGATATTTTGAGACCACTG	1233
QY	436	GGTGGCTGACAGATCATCAACCCATTGGTGGAGATGATGACACTTTGAGCAAAATCA	495
Db	1234	GATTGTGCACAGAAATTTCAGAGTGTCCCTGTTGGCTGTGGATGATGCACAGAACT	1293
QY	496	GCTCATATGACCGCAACTTGCAGGTGTCCCTCTCATCGTGGACGAAATGTACGAAACT	555
Db	1294	GCCCTCGATGAGAGCCCGGACATGTACTGGAATTAACCCGAGGCCACAGCCCCCTTACAC---	1356
QY	556	TCCCCCGCTGAGAGGACAGACTGTGGAGATGAGAGCACAGCCGACCACTTACACTGT	615
Db	1351	AGCTGCTTCCGCCGCCAGTTCGCTCCAGGCTCTCTTTATGGGCTCCACTTCAACATCAGCCT	1410
QY	616	GGCCACGGCGGCGGAGTCTCTGCGGCGCTCATTTCTGGGCTCCCACTTTCAACCTGGCAT	675
Db	1411	GAAACAAAGAGAGATGAGTTCCAGC	1436
QY	676	GAGCGACGACCTTGAGCAGAGCCTTGC	701

```

RESULT 13
US-10-198-846-11346
; Sequence 11346, Application US/10198846
; Publication No. US2003099974A1
; GENERAL INFORMATION:
; APPLICANT: Liilie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinhmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11346
; LENGTH: 853
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
; LOCATION: 43, 50, 51, 54, 57, 853
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346

```

Query Match	4.8%	Score 116;	DB 14;	Length 853;
Best Local Similarity	68.2%;	Pred. No. 2.8e-21;		
Matches 161; Conservative	0;	Mismatches 75;	Indels 0;	Gaps 0;

Qy	350	GGGCTTCTAAGTGAAGCGCTGGTGGTGAACCGCGGTGGAAACAGTACAGAAACCTGGCCGTG	409
Db	272	GGGGTTTTATGTTACTCTGGTAGTAAACGATGGTGGACCAAGTTTGAAATTTGGCCCTG	331
Qy	410	GCCCAACCCCTCATAGAGCTGGTGGGGCTTCGTGTCMAAGGCACGACGACAGGCCG	468
Db	332	GCCAGACAGGCTATATGTTTCTCATCTCTAGCAGGTTCACGGAAGCAGACAGCACGGGGC	391
Qy	470	GCTGCTGCGGCGCAGCTTCATCCGGTACGCCAACCTGGGCAACGTGCTCATCTGGCGAG	529
Db	392	CTGCTTTAAGAGACCGCTGATGCGCTACGTCAATCTCACTCCCTGCTCATCTTTGCTC	451
Qy	530	CGTCAGCACCGCAGTCTACAAAGCGTTTCCCGACGGCCACAGACTGTGGTCAAGAG	585
Db	452	GGTGAACACTGCTGTGTACAAAGATTTCCACATAGAACACCGCTGGTTGAGAG	507

RESULT 14

```

US-10-198-846-2561/c
; Sequence 2561, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2561
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 390..462
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-2561

```

```

Query Match      4.3%; Score 104; DB 14; Length 462;
Best Local Similarity 67.6%; Pred. No. 4.2e-18;
Matches 146; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

```

```

QY 349 TGGGCTTACAGTACGCTGCTGCTGACCCGCTGCTGGAACCACTACAGAACCTGCGCT 408
DB 255 TAGGTTTATGTTACTCTGTTAGTGAACCATGTGGAACAGTTTGTGATTTGCTT 196
QY 409 GGGCCGACCGCTCATGAGCTGTGTGGGCTTCTGCAAGCAAGACGAGCAAGCC 468
DB 195 GGGCAGACAGGCTATGTTCTCTCATCTTACAGTGTTCAGGAGCAAGACGAGCC 136
QY 469 GCGTCTGCGGCGGCGCTGATCCGCTACGCAACTGGCAAGTGTCTCATCTTGGCA 528
DB 135 GCGTCTTGAAGAGAGCGCTGATGCGCTACGTCATCTACCTCCGCTGCTATTTGCT 76
QY 529 GCGTCAGCACCGCACTGTACAGGCGCTTCCCGACGG 564
DB 75 CGGTGAGCACTGCTGTACTCTCGGCGCGCACCGCG 40

```

```

RESULT 15
US-10-027-632-174961/c
; Sequence 174961, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002

```

```

; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 174961
; LENGTH: 3186778
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(3186778)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-174961

```

```

Query Match      4.3%; Score 103.6; DB 15; Length 3186778;
Best Local Similarity 77.5%; Pred. No. 5.8e-16;
Matches 138; Conservative 0; Mismatches 39; Indels 1; Gaps 1;

```

```

QY 418 GCGTCATGAGCTGTGTGTGCGGCTTGTGCAAGCAAGACGAGCAAGCGGCTGCTGC 477
DB 285842 GCTCTGAGAGCTGTGTGTC-OGCTCGCGAGGCAAGAGAGAGCAAGCGGCTGAAG 285784
QY 478 GGGCAGGCTCATCGCTACGCCAACCTGAGCAAGTCTCATCTGCGAGGCTACAGCA 537
DB 285783 GAGCATGCTCTCTTCCGCTTGCCCAACGCGCAACGTCATTAACTGCGGAAAGCGCAGCA 285724
QY 538 CCGCATCTACAAAGCGCTTCCCGAGCGCCCAAGCACTGTGTGCAAGAGGCTTATGAC 595
DB 285723 CCTCGAACAAAGCGCTTCCCGAGCGCGGCAAGCAAGGTGCAAGAGATGTGAC 285666

```

```

Search completed: March 26, 2004, 11:54:40
Job time : 815.507 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2004, 22:13:32 ; Search time 134.65 Seconds
(without alignments)
7896.639 Million cell updates/sec

Title: US-09-622-964-28

Perfect score: 1916
Sequence: 1 GTGCCAAGCAGCAGTACATCA.....aaaaaaaaaaaaaaaaa 1916

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.2	2.7	1361	4	US-09-489-847-64
2	51.8	2.7	1931	3	US-09-019-942-2
3	51.8	2.7	1931	4	US-09-019-942-2
4	51.8	2.7	1931	4	US-09-099-041A-1
5	51.8	2.7	1931	4	US-09-245-281-1
6	51.8	2.7	1931	4	US-09-470-271-2
7	51.8	2.7	1931	4	US-09-207-359B-1
8	51.8	2.7	1931	4	US-09-340-620A-1
9	51.8	2.7	1931	4	US-09-865-364-1
10	51.4	2.7	1931	4	US-09-748-537-2
11	51.4	2.7	2619	4	US-08-983-502-17
12	51.4	2.7	2619	4	US-09-516-747-17
13	51.4	2.7	2619	5	PCT-US86-10521-17
14	51.4	2.7	2887	4	US-08-983-502-14
15	51.4	2.7	2887	4	US-09-516-747-14
16	51.4	2.7	2887	5	PCT-US96-10521-14
17	51.4	2.7	2060	4	US-09-345-473B-5
18	50.8	2.7	1446	4	US-09-593-359B-1
19	50.6	2.6	1582	3	US-08-545-196B-10
20	48.6	2.6	1582	3	US-08-545-196B-12
21	48.4	2.5	1584	4	US-09-807-258-11
22	47.8	2.5	2053	4	US-09-023-655-854
23	47.6	2.5	1046	1	US-08-361-467B-4
24	47.6	2.5	1046	1	US-08-484-332C-4
25	47.4	2.5	458	1	US-08-524-757-1
26	47.4	2.5	1493	1	US-08-340-820-24
27	47.4	2.5	1493	1	US-08-593-535-24

28	47.2	2.5	780	2	US-08-540-804-36	Sequence 36, Appl
29	47.2	2.5	780	3	US-08-590-399-36	Sequence 36, Appl
30	47.2	2.5	1952	1	US-08-333-359-1	Sequence 1, Appl
31	47.2	2.5	1952	1	US-08-463-694-1	Sequence 1, Appl
32	47.2	2.5	1952	1	US-08-694-501-1	Sequence 1, Appl
33	47.2	2.5	3214	1	US-08-484-105-17	Sequence 1, Appl
34	47.2	2.5	3214	1	US-08-484-106-17	Sequence 1, Appl
35	47.2	2.5	3722	4	US-10-164-595-9	Sequence 17, Appl
36	47.2	2.5	3862	4	US-10-164-595-5	Sequence 5, Appl
37	47.2	2.5	3937	4	US-10-164-595-7	Sequence 7, Appl
38	47.2	2.5	3985	4	US-10-164-595-3	Sequence 3, Appl
39	47.2	2.5	4895	3	US-09-053-866-1	Sequence 1, Appl
40	47.2	2.5	4895	4	US-09-479-130-1	Sequence 1, Appl
41	47.2	2.5	4895	4	US-09-472-130A-1	Sequence 1, Appl
42	47.2	2.5	467	2	US-08-841-349-18	Sequence 18, Appl
43	47.2	2.5	467	4	US-09-431-184A-18	Sequence 18, Appl
44	47.2	2.5	1313	4	US-09-149-476-112	Sequence 112, App
45	47.2	2.5	1835	4	US-09-564-808-5	Sequence 5, Appl

ALIGNMENTS

```
RESULT 1
US-09-489-847-64
; Sequence 64, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT APPLICATION NUMBER: US/09/489,847
; EARLIER FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/117130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 64
; LENGTH: 1361
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-489-847-64

Query Match      2.7%; Score 52.2; DB 4; Length 1361;
Best local similarity 71.1%; Pred. No. 0.00069;
Matches 69; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY      1820 CTGATTTTACATCTTCCCAACTAGAGCTTATATACGTGATATCTTTAGGT 1879
      |||||
DB      1216 CAGTTGTGAAAATATATATTAGCAACTCATGAGATTAACCCATTAATTATTTATTA 1275
      |||||

QY      1880 GAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
      |||||
DB      1276 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1312
      |||||

RESULT 2
US-09-019-942-2
; Sequence 2, Application US/09019942
; Patent No. 6033855
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
```

```

; TITLE OF INVENTION: DOMAIN POLYPEPTIDES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/019,942
; FILING DATE: 06-FEB-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Weiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/068001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1931 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-09-019-942-2
;
Query Match
Best Local Similarity 2.7%; Score 51.8; DB 3; Length 1931;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
;
QY 1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTATTAAGCGTAAATATCTT 1873
DB 1825 AAAAGCATGTAAAGTGAAGTCTTTTCAAGAAATGTGTTTCAATAAGGATTTTATTA 1884
;
QY 1874 TTAGGTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1916
DB 1885 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1927
;
RESULT 3
US-09-099-041A-1
; Sequence 1, Application US/09099041A
; Patent No. 6340576
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
; TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
; FILE REFERENCE: 07334-076001
; CURRENT APPLICATION NUMBER: US/09/099,041A
; CURRENT FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 09/019,942
; PRIOR FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (214)...(1833)
;
US-09-099-041A-1
```

```

Query Match
Best Local Similarity 2.7%; Score 51.8; DB 4; Length 1931;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
;
QY 1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTATTAAGCGTAAATATCTT 1873
DB 1825 AAAAGCATGTAAAGTGAAGTCTTTTCAAGAAATGTGTTTCAATAAGGATTTTATTA 1884
;
QY 1874 TTAGGTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1916
DB 1885 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1927
;
RESULT 4
US-09-245-281-1
; Sequence 1, Application US/09245281
; Patent No. 6369196
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 07334/118001
; CURRENT APPLICATION NUMBER: US/09/245,281
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: US 09/207,359
; EARLIER FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: US 09/099,041
; EARLIER FILING DATE: 1998-06-17
; EARLIER APPLICATION NUMBER: US 09/019,942
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (214)...(1833)
;
US-09-245-281-1
;
Query Match
Best Local Similarity 2.7%; Score 51.8; DB 4; Length 1931;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
;
QY 1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTATTAAGCGTAAATATCTT 1873
DB 1825 AAAAGCATGTAAAGTGAAGTCTTTTCAAGAAATGTGTTTCAATAAGGATTTTATTA 1884
;
QY 1874 TTAGGTGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1916
DB 1885 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1927
;
RESULT 5
US-09-470-271-2
; Sequence 2, Application US/09470271
; Patent No. 6410689
; GENERAL INFORMATION:
; APPLICANT: Bertin, John
; TITLE OF INVENTION: GENES ENCODING CASPASE RECRUITMENT
; TITLE OF INVENTION: DOMAIN POLYPEPTIDES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
```



```
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/470,271
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/019,942
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/068001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1931 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-470-271-2
```

```
Query Match
Best Local Similarity 68.9%; Score 51.8; DB 4; Length 1931;
Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
QY 1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTAACTGAATATTTCTT 1873
DB 1825 AAAAGCATGTAACTGATCTGTTTTCAGAAAGAAATGTGTTCTAAAGATATTATTA 1884
QY 1874 TTAGTGAAAAA 1916
DB 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927
```

```
RESULT 6
US-09-207-359B-1
Sequence 1, Application US/09207359B
Patent No. 6469140
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
FILE REFERENCE: 07334-112001
CURRENT APPLICATION NUMBER: US/09/207,359B
CURRENT FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (214)...(1833)
US-09-207-359B-1
```

```
Query Match
Best Local Similarity 68.9%; Score 51.8; DB 4; Length 1931;
Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
QY 1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTAACTGAATATTTCTT 1873
DB 1825 AAAAGCATGTAACTGATCTGTTTTCAGAAAGAAATGTGTTCTAAAGATATTATTA 1884
```

```
QY 1874 TTAGTGAAAAA 1916
DB 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927
```

```
RESULT 7
US-09-340-620A-1
Sequence 1, Application US/09340620A
Patent No. 6482933
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THEREOF
FILE REFERENCE: 07334-124001
CURRENT APPLICATION NUMBER: US/09/340,620A
CURRENT FILING DATE: 1999-06-28
PRIOR APPLICATION NUMBER: US 09/245,281
PRIOR FILING DATE: 1999-02-05
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 71
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (214)...(1833)
US-09-340-620A-1
```

```
Query Match
Best Local Similarity 68.9%; Score 51.8; DB 4; Length 1931;
Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
```

```
QY 1814 AAACACCTGATTTTACGATCTTTCCCAACTAAGAGTTTAACTGAATATTTCTT 1873
DB 1825 AAAAGCATGTAACTGATCTGTTTTCAGAAAGAAATGTGTTCTAAAGATATTATTA 1884
QY 1874 TTAGTGAAAAA 1916
DB 1885 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1927
```

```
RESULT 8
US-09-865-364-1
Sequence 1, Application US/09865364
Patent No. 6613521
GENERAL INFORMATION:
APPLICANT: Bertin, John
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED
FILE REFERENCE: 07334-112001
CURRENT APPLICATION NUMBER: US/09/865,364
CURRENT FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 09/207,359
PRIOR FILING DATE: 1998-12-08
PRIOR APPLICATION NUMBER: US 09/099,041
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: US 09/019,942
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1931
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (214)...(1833)
```

US-09-865-364-1

Query Match 2.7%; Score 51.8; DB 4; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1814 AACACCTGATTTTACGATCTTCCCAACTAGAGATTATTAACGTGAATTTCTT 1873
DB 1825 AAAAGCATGTAGTCTGTTTTCAGAGAAATGTGTTTCATTAAGATATTTATTA 1884
QY 1874 TTAGTGAGAAAAA 1916
DB 1885 AA 1927

RESULT 9
US-09-748-537-2
Sequence 2, Application US/09748537
Patent No. 6680167
GENERAL INFORMATION:
APPLICANT: Bertin, John
APPLICANT: Chao, Moses V.
TITLE OF INVENTION: NOVEL MOLECULES OF THE CARD-RELATED PROTEIN FAMILY AND USES THERE
FILE REFERENCE: 07334-316001
CURRENT FILING DATE: 2000-12-26
PRIOR FILING DATE: 1998-06-17
PRIOR FILING DATE: 1998-02-06
NUMBER OF SEQ ID NOS: 14
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1931
TYPE: DNA
ORGANISM: Homo sapiens
US-09-748-537-2

Query Match 2.7%; Score 51.8; DB 4; Length 1931;
Best Local Similarity 68.9%; Pred. No. 0.001;
Matches 71; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 1814 AACACCTGATTTTACGATCTTCCCAACTAGAGATTATTAACGTGAATTTCTT 1873
DB 1825 AAAAGCATGTAGTCTGTTTTCAGAGAAATGTGTTTCATTAAGATATTTATTA 1884
QY 1874 TTAGTGAGAAAAA 1916
DB 1885 AA 1927

RESULT 10
US-08-983-502-17
Sequence 17, Application US/08983502
Patent No. 639327
GENERAL INFORMATION:
APPLICANT: David WALLACH
APPLICANT: Mark P. BOLDIN
APPLICANT: Tanya M. GONCHAROV
APPLICANT: Yuri V. GOLITSEV
TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
TITLE OF INVENTION: AND OTHER PROTEINS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/983,502

FILING DATE: 16-JAN-1998

PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10521

FILING DATE: 14-JUN-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 114,615

FILING DATE: 16-JUL-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 114,986

FILING DATE: 17-AUG-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 115,319

FILING DATE: 14-SEP-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 116,588

FILING DATE: 27-DEC-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117,932

FILING DATE: 16-APR-1996

ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH-19
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2619 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-983-502-17

Query Match 2.7%; Score 51.4; DB 4; Length 2619;
Best Local Similarity 72.0%; Pred. No. 0.0015;
Matches 67; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1824 ATTTAGCATCTTCTCCCAACTAGAGATTATTAACGTGAATTTCTT 1883
DB 2527 ATTTAGCATCTTCTTCCCAAACTATTTTATGAATTAATATGATCTTAA 2586
QY 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
DB 2587 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2619

RESULT 11
US-09-516-747-17
Sequence 17, Application US/09516747
Patent No. 6586571
GENERAL INFORMATION:
APPLICANT: David WALLACH
APPLICANT: Mark P. BOLDIN
APPLICANT: Tanya M. GONCHAROV
APPLICANT: Yuri V. GOLITSEV
TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
TITLE OF INVENTION: AND OTHER PROTEINS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

```
/
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/516,747
/ FILING DATE: 01-Mar-2000
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/983,502
/ FILING DATE: <Unknown>
/ APPLICATION NUMBER: IL 114,615
/ FILING DATE: 16-JUL-1995
/ APPLICATION NUMBER: IL 114,986
/ FILING DATE: 17-AUG-1995
/ APPLICATION NUMBER: IL 115,319
/ FILING DATE: 14-SEP-1995
/ APPLICATION NUMBER: IL 116,588
/ FILING DATE: 27-DEC-1995
/ APPLICATION NUMBER: IL 117,932
/ FILING DATE: 16-APR-1996
/
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Browdy, Roger L.
/ REGISTRATION NUMBER: 25,618
/ REFERENCE/DOCKET NUMBER: WALLACH=19
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 628-5197
/ TELEFAX: (202) 737-3528
/
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2619 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ MOLECULE TYPE: cDNA
/
/ SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-516-747-17

Query Match          2.7%; Score 51.4; DB 4; Length 2619;
Best Local Similarity 72.0%; Pred. No. 0.0015;
Matches 67; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1824 ATTTAGCATCTTCCCAACTAGAGTTTATTAACGTGAATATCTTTAGTGAAA 1883
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2527 ATTTAGTATCTTTTCCAAACATTTTATAGATAAATATTAATTGATCTTAAA 2586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2587 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2619
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 12
PCT-US96-10521-17
/ Sequence 17, Application PC/TUS9610521
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
/ TITLE OF INVENTION: AND OTHER PROTEINS
/ NUMBER OF SEQUENCES: 34
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/10521
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 114,615
/ FILING DATE: 16-JUL-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 114,986
/ FILING DATE: 17-AUG-1995
/ PRIOR APPLICATION DATA:
```

```
/
/ APPLICATION NUMBER: IL 115,319
/ FILING DATE: 14-SEP-1995
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 116,588
/ FILING DATE: 27-DEC-1995
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 117,932
/ FILING DATE: 16-APR-1996
/
/ INFORMATION FOR SEQ ID NO: 17:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2619 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/
/ MOLECULE TYPE: cDNA
/
/ PCT-US96-10521-17

Query Match          2.7%; Score 51.4; DB 5; Length 2619;
Best Local Similarity 72.0%; Pred. No. 0.0015;
Matches 67; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

QY 1824 ATTTAGCATCTTCCCAACTAGAGTTTATTAACGTGAATATCTTTAGTGAAA 1883
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2527 ATTTAGTATCTTTTCCAAACATTTTATAGATAAATATTAATTGATCTTAAA 2586
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2587 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2619
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-08-983-502-14
/ Sequence 14, Application US/08983502
/ Patent No. 6399327
/ GENERAL INFORMATION:
/ APPLICANT: David WALLACH
/ APPLICANT: Mark P. BOLDIN
/ APPLICANT: Tanya M. GONCHAROV
/ APPLICANT: Yuri V. GOLTSSEV
/ TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
/ TITLE OF INVENTION: AND OTHER PROTEINS
/ NUMBER OF SEQUENCES: 34
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Browdy and Neimark
/ STREET: 419 Seventh Street N.W., Ste. 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20004
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/983,502
/ FILING DATE: 16-JAN-1998
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/10521
/ FILING DATE: 14-JUN-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 114,615
/ FILING DATE: 16-JUL-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 114,986
/ FILING DATE: 17-AUG-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 115,319
/ FILING DATE: 14-SEP-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: IL 116,588
/ FILING DATE: 27-DEC-1995
/ PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: IL 117,932
FILING DATE: 16-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=19
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2887 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-983-502-14

Query Match 2.7%; Score 51.4; DB 4; Length 2887;
Best Local Similarity 72.0%; Pred. No. 0.0016;
Matches 67; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1824 ATTTAGCATCTTTCCCAACTAAGAGTTAATAACGTGAATTCCTTTAGGTAAA 1883
Db 2791 ATTTAGTATCTTTTCCAAACATTTTAAATGATATAATATATGATCTTAAA 2850
Qy 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
Db 2851 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2883

RESULT 14

US-09-516-747-14
Sequence 14, Application US/09516747
Patent No. 6586571
GENERAL INFORMATION:
APPLICANT: David WALLACH
Mark P. BOLDIN
Tanya M. GONCHAROV
Yury V. GOLITSEV
TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street N.W., Ste. 300
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/516,747
FILING DATE: 01-Mar-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/983,502
FILING DATE: <Unknown>
APPLICATION NUMBER: IL 114,615
FILING DATE: 16-JUL-1995
APPLICATION NUMBER: IL 114,986
FILING DATE: 17-AUG-1995
APPLICATION NUMBER: IL 115,319
FILING DATE: 14-SEP-1995
APPLICATION NUMBER: IL 116,588
FILING DATE: 27-DEC-1995
APPLICATION NUMBER: IL 117,932
FILING DATE: 16-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Browdy, Roger L.

REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH=19
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 628-5197
TELEFAX: (202) 737-3528
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2887 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-516-747-14

Query Match 2.7%; Score 51.4; DB 5; Length 2887;
Best Local Similarity 72.0%; Pred. No. 0.0016;
Matches 67; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1824 ATTTAGCATCTTTCCCAACTAAGAGTTAATAACGTGAATTCCTTTAGGTAAA 1883
Db 2791 ATTTAGTATCTTTTCCAAACATTTTAAATGATATAATATATGATCTTAAA 2850
Qy 1884 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1916
Db 2851 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2883

RESULT 15

PCT-US96-10521-14
Sequence 14, Application PC/TUS9610521
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: MODULATORS OF THE FUNCTION OF FAS RECEPTORS
NUMBER OF SEQUENCES: 34
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/10521
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 114,615
FILING DATE: 16-JUL-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 114,986
FILING DATE: 17-AUG-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 115,319
FILING DATE: 14-SEP-1995
APPLICATION NUMBER: IL 116,588
FILING DATE: 27-DEC-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 117,932
FILING DATE: 16-APR-1996
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2887 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
PCT-US96-10521-14

Query Match 2.7%; Score 51.4; DB 5; Length 2887;
Best Local Similarity 72.0%; Pred. No. 0.0016;
Matches 67; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy	1824	ATTTACGATCTTTCCCAACTAAGAGTTTAATAACGGAATATTCCTTTAGTGAAA	1883
Db	2791	ATTTAAGTATCTTTTCCCAAAACATTTTAATAGATATAATTAATTGATCTTAA	2850
Qy	1884	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	1916
Db	2851	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	2883

Search completed: March 26, 2004, 05:11:13
 Job time : 136.65 secs

This page blank (uspio)

OY	251	GTCTGGGTTCTAATGTTACATATGGTGTGGCCGCTGCTGAGCCAGTACAGAACTTG	310
Db	469	GTCTGGGTTTAAATGTTACTCTGTGTAATGAACCGAATGTGGAAACGATTGTGAATTGG	528
OY	311	CCGTGGCCCGACCGCCTCATGATATCCAGGTGTCTAGCTTCTGTGAGGGCAGAAATGAGAA	370
Db	529	CCCTGGCCAGACAGGCTATATGTCTCTCATCTCTAGCAGTGTCTACGGAAAGGAGACAC	588
OY	371	GGCCGTTTGTGCGGGCGACGCTATCCGTAACGCAATCCGAGGGCCAAAGTCTCATCTG	430
Db	589	GGGGCGCTGCTTGAAGAGACGCTGATGGGCTACGTCAATCTCACTCCCTGCTATCTTT	648
OY	431	CGCAGCATACAGCACCCTCGCTCTAAGCCATTAAGCCCTTCCACTCTTCAACAATGCTGAC	490
Db	649	CGCTCGGTGACACATGCTGTGTACAAAAGATTTCCCAATATGACACACGTGTGAAGCA	708
OY	491	GCTTTTATGACCCATGGGGAACTATAGACATGTGCAGAAAGTTGGGCTTACCAACAACA	550
Db	709	GGTTTATGACAAACGATGAAAGGAAATATTAACCAACCAAGTCTCCATCTGAA	768
OY	551	TTCTGGGGGCCCTGGGTGTGGTTTGGCCAACTGTTCATATGAAGAGCTTACTTGGAGTGA	610
Db	769	TATTTGGTTCCATTCATCTGTGTTTGAATCTTGCACTTAAGCCGGAAATGAAGTAGA	828
OY	611	ATCCGGGACACCGTCTGCTCCAGAGCTGATGAATGAAGTGTGTACTTGGCTATCTGAC	670
Db	829	ATCAGAGACAGTGTGATCTGCATCATTTGATATGACTGAATGAATGCATACCGCTCTTGG	888
OY	671	TGTGAGACAGCTGTATGCTTAGACCTGAGACTGGAATGATCCATTTGGATGACACAGGTGTG	730
Db	889	TGCACCTCTTATTCGGTTATGACTGAGGTGGGATTCGGCTGTATTACACCCAGTTGTC	948
OY	731	AACGTGGCAGTATACAGCTTTTTCCTTGCATGCTTGATATGGGAGGACGTTTCTGAACCA	790
Db	949	ACTCTGTGCTCTAATACCTCTTCTTTCGCTGACGATTTGACACGCAATTTTGTGATCCC	1008
OY	791	AACAAGACCTACCCAGGCCATGAGATGATGCTGTGACGCTTACCAATCCTGCA	850
Db	1009	ACCAAGGCTACGCAAGGCGATGACTTGGATCTTTCATTTCCATTTCAACCTCTCTACAA	1068
OY	851	TTCTATATCTCATATGGGCTGGCTGAGGTGGCAGAAACGCTCATCAACCCCTTGGAGGAG	910
Db	1069	TTCTTCTTCTATGACAGATGGCTTAAAGTAGACAGACGCTTATCAACCTTTTGGAGAA	1128
OY	911	GACGATGATGATTTTGAAGCTATCTGATCATTTGACAGAAACCTGACAGTGTCTCTGTG	970
Db	1129	GATGATGATGATTTTGAACATTAACCTGGTGCATTTGACAAAAATTTGCAGAGTCTCTTTTA	1188
OY	971	TCCGAGATGGGATGACACAGAACTTGCTCCCATGAGAACGTGACATCTACTGAAACGAG	1030
Db	1189	GCTGTGAGCAAAATGACATATAGCTTTACCAAGATGAAAGAAAGACATTTACTGGAGCAT	1248
OY	1031	GCAGGCGCTCAGCGGCCCTACACAGACTGTTCTGCAAGTCTCGCGGCAATTCCTTCAATG	1090
Db	1249	TCTGTGCTCGCCCAACATACATATGGCAGCTGTGTACTGTCATACCTCATATTTCTG	1308
OY	1091	GGCTCACCTTACATACAGCCT	1113
Db	1309	GGGTCAACGTCCAAATGGGGCT	1331

RESULT 2
US-09-746-783-3
; Sequence 3, Application US/09746783
; Publication No. US20030044935A1
; GENERAL INFORMATION:

Jacobs, Kenneth
McCoy, John M.
Lavallie, Edward R.
Racie, Lisa A.
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.

Howes, Steven H.
Rechtel, Kim
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM
NUMBER OF SEQUENCES: 231
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Milasincic, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Query Match 11.9%; Score 227.6; DB 10; Length 1263;
Best Local Similarity 67.7%; Pred. No. 2,66-53;
Matches 383; Conservative 0; Mismatches 144; Indels 39; Gaps 3;

Query Match	Similarity	11.9%	Score 227.6;	DB 10;	Length 1263;
Best Local	Similarity	67.7%	Pred. No.2.6e-53;		
Matches	383;	Conservative	0;	Mismatches 144;	Indels 39; Gaps 3;
QY	884	GACAGCTCATCAACCCCTTGCGGAGAGACGATGATGATTTTGGACTTAATCGATCATTT	943		
Db	1	GAGCAGCTCATCAACCCCTTTGGAGAGATGATGATGATTTTGGAGCCAACTGGATTGCTC	60		
QY	944	GACAGAAACCTGCAGGTGTCCCTTGTGCTCGTGATGGATGCACAGAACTTGTCTCCC	1003		
Db	61	GACAGGAATTTGCAGGTGTCCCTGTGTGGCTGTGATGATGATGCACAGAACTTGTCTCGG	120		
QY	1004	ATGGAACCTGACATGTACTGGAACGAGGACAGCCTCAGCCGCCCTACAGAGCTGCTTCT	1063		
Db	121	ATGGAAGCCGAGACATGTACTGTGAATAAGCCAGGACAGCCCTTACAGAGCTGCTTCC	180		
QY	1064	GCCAGGTGTGCGCGGCAATTCCTTATATGGGCTCCACCTTCAACATCAGCCTTAAGAAAGAA	11223		
Db	181	GCCAGTGTCCGTGAGCCTCCTTTATGAGCTTCCACCTTCAACATCAGCCTTAAGAAAGAG	240		
QY	1124	GACTTAGAGCTTTGGTCAAAAAGAGGAGGCTGACACGATTAAGAAAGAGGTGGCTATTAGC	1183		
Db	241	GAGATGAGTTCCAGCCCAATCAGAGG-----ACGAGGAAGATGTCTACCGCT	288		
QY	1184	AGCACCATAGGCTGTCTTTAGGACTGCACTCAACCCCAAAAATAACATCTTCCCTTGAAGAAC	124		
Db	289	GGCATCATTTGGCCGCTTCTCTAGGCTCGTAGCTTCATGATCAACATCTCTCCAGGGCAAAAC	348		
QY	1244	TTAAAGACCAACTATTGTGTCTTAAGAA-----CCCCCTCTCGAAGGCCAAGTGAAG	129		
Db	349	TCAAGACCAAACTACTGTGCGCCAAAGGGAATCTCTTCTCCAGAGAGGCTGTGCCAAA	408		
QY	1298	GATGCGAACCAAGAAAACCAAGAAAT-----GTCTGAAATTT	1336		
Db	409	AACCAACAAGCAGCCCAACAGAACCTTATGGGCCCAAGAAACACAAAGCCTGTGAAGCTT	468		

Qy 1337 AAGGCTGTGACTTCTTGAATGTGTCCAGAGTTTAAGAGAGAGCTCCATTGTGC 1396
Db 469 AAGGCTGTGAGACGCTTCAAGTCTGGCCCACTGTATCAGAGGCGACGCTACTACAGTGCC 528
Qy 1397 CCACAGGACCCGACGACCCACCTTAC 1422
Db 529 CCACAGACGCGCTCTCAGCCCACTCC 554

RESULT 3

US-09-814-353-19504/c
; Sequence 19504, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Little, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19504
; LENGTH: 620
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19504

Query Match 10.1%; Score 194.4; DB 10; Length 620;
Best Local Similarity 63.9%; Pred. No. 4.1e-44;
Matches 294; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

Qy 9 CCATGACTATACCTACACAAAGTACGCAATCCCGCTCGGTTGTTCTGCTCC 68
Db 464 CCATGACGGTTTCACTACCTCAAGTGGCGGAGCGCTTCGAGTTCTGCGCC 405
Qy 69 TCCTCTGCTGGGAGGAGCATCTAGAGCTGTGTATGAGAAATTCCTGTTCTTA 128
Db 404 TCTCTCCGCTGAGGAGGAGATTAAGCTCTCTACAAAGAAATTCCTCTTTGG 345
Qy 129 TATTCCTACTATTCATCCGTGAGCTTACAGAAATGTTCTCTGATGATCAGAC 188
Db 344 GGGCCCTGTAGCGTGTAGATCACTACCGGCTGTCTGATGATCAGAGAGGT 285
Qy 189 TGTGTTTGAAGAGCTGCTGTGATGAGACGACTACATTACGCTCATATCTCT 248
Db 284 AGGTGATCTCAGGTGGCCCGGTACTGCAACGCTCAGACGCTCATTCCTTGTCT 225
Qy 249 TCGTTCTGGGTTTCTATGTTAATTGATGAGTGGAGCGGTGAGACGATCGAAGT 308
Db 224 TTGTATTTGGGTTTCTATGATCTCTGTGTGAGACGCTGGTGGTCCAGTACAGAGA 165
Qy 309 TCCCGTGGCCGACCGCTCATGATCAGGTGTCTTCTGTGAGAGGCAAGATGAG 368

Db 164 TCCCGTGGCAGACGACTGATGTGCTCATCTCGGCTAGCGGTGACCGCGTGACAC 105
Qy 369 AAGCGGTTTGTGCGGGCGGACGCTCATCGGCTAGGCGCATTCGAGGCGCAAGTCTATCC 428
Db 104 GGGCGCGCTGTGCGCGGACCTCATCGGCTAGCGGAACTGGCGTTGTGTGTGTGC 45
Qy 429 TCGCAGCATCAGACCTGCTGTACAGAGCGTTTCCAC 468
Db 44 TCGCTCGGTGACGACCGCGGTCTTAAGCGCTTCCCA 5

RESULT 4

US-09-768-826-16
; Sequence 16, Application US/09768826
; Patent No. US20020012966A1
; GENERAL INFORMATION:
; APPLICANT: Shi et al.
; TITLE OF INVENTION: 18 human secreted proteins
; FILE REFERENCE: PFS12P1
; CURRENT APPLICATION NUMBER: US/09/768,826
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: PCT/US00/22350
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/148,759
; PRIOR FILING DATE: 1999-08-16
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (1135)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1148)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1166)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1174)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1181)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1209)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1229)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1266)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1285)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1287)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1290)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1295)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1305)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (1324)

```
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1339)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1341)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1343)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1345)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (1348)
OTHER INFORMATION: n equals a,t,g, or c
US-09-768-826-16
```

```
Query Match 9.8%; Score 187.8; DB 9; Length 1350;
Best Local Similarity 60.1%; Pred. No. 4,8e-42; Indels 48; Gaps 2;
Matches 377; Conservative 0; Mismatches 202;
```

```
544 CAACATCTGTGGGTCCTGGGTGTGTTGCAACTGTGCAATGAAGCCATCTTGG 603
58 CAACAGTACTGGGTCCCTGCTGCTGTTCAACCACTGGCGCCGACGCGAGGGA 117
604 AGGTGAATCCGGGACACCGTCTGCTGCAAGCTGATGATGAGGTGTACTTTGG 663
118 CGGGCAATACGTGACGATATCGCTCTGCTACTTTTGAAGACTCAACAGTACCG 177
664 TACTAGTGTGACAGCTGTATGCCACGATGAAATGCCATGTTGTGACACA 723
178 AGCCAGTGCAGATGCTATTCATCATATACGTGATCAGATCCCTGCTTACACCA 237
724 GGTGAGCAGTGTGACAGTATACAGCTTTTCTTGATGCTTGATGGAGGAGCTTCT 783
238 AGTGTGACATAGCCGTCTACTCTTTTGGCCCTCCCTGTTGGCGCAGTTTGT 297
784 GAACCAACAAAGACTACCCAGGCGCATGATGATGATGATGATGATGATGATGAT 826
298 GGAGCGAGAGGAGGAGGCGGCTGCCAACCCTCAAGACCTTCTGAGCCAGGCGAG 357
827 -----GTGCTGTCTTCAACATCTGCAATTTCTATT 858
358 CCCAGCCCTGGGAGACCCGGAATGTACGTGCTCTCACCACCTCTGCTGAGTTCTT 417
859 CTACATAGGCTGGCTGAAGGTGGCAGAACAGCTCATCAACCCCTTGGGGAGGACGAT 918
418 CTATGTGCTGGCTCAAGGTGGCTGACAGATCATCAACCCATTTGGTGAAGATGAT 477
919 TGATTTTGAAGACTACTGATCTTGAACAGAACTGTGAGGTCTCTGTTTCTGCTGA 978
478 CGACTTTGAGACAAATCAGCTCATAGACGCAACCTTGAGGTCTCTGCTATCCGTGA 537
979 TGGAGTGCACCGAAGCTTGTCTCCAGTGAAGTGAATGATGATGATGATGATGATG 1038
538 CGAAATGTACAGAACTTCTCCCGCTGAGAAAGACACTGAGTGAAGGACAGC 557
1039 TCAGCGCCCTTACAC--AGCTGCTTCTGCGAGGTCTCGCGGCAATTTCTTCA 1095
598 GAGGCAACCTTACACTGTGCGACAGGCGGCGGAGGTCTCTGCGGCGCTCAATCT 657
1096 CACCTTCAACATCAGCTTAAAGAA 1122
658 CACCTTCAACCTCGCATGAGCGACGA 684
```

```
RESULT 5
US-10-198-846-11070
; Sequence 11070, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Millie, James
```

```
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
CURRENT FILING DATE: 2002-07-18
PRIOR APPLICATION NUMBER: 60/306,220
PRIOR FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11070
LENGTH: 1292
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1287-1288, 1289, 1290, 1291, 1292
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11070
```

```
Query Match 9.5%; Score 181.2; DB 14; Length 1292;
Best Local Similarity 61.4%; Pred. No. 3,4e-40; Indels 0; Gaps 0;
Matches 291; Conservative 0; Mismatches 183;
```

```
255 TGGGTTTCTATTTTCACTTTGTTGTTGAGCCGCTGGTGGAGCCAGTACAGAACTTGGCGT 314
318 TAGGGTTTATTTTCTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 377
315 GGGCCGACCGGCTCAATGATCCAGGTGTCTAGCTTGTGTTGTTGTTGTTGTTGTTGTT 374
378 GGCACAGACGGTTAATGTTCTTCTATCTTACAGATGTTTCAACGAGGACGACGAGGCG 437
375 GTTGTGCGGGGACGCTCATTCGCTACGCTATCCCTGGGCGCAAGTCTCATCTGCGCA 434
438 GCTGTGTTGAGAGAGCGCTGATGCGCTACGTAATCTCACTCCCTGCTCATCTTTCGT 497
435 GCATGACGACCTCGGTCTTCAAGCGCTTCCACTCTTACCAACCTGTTGTTGTTGTTGTT 494
498 CGGTGAGCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 557
495 TTATGACCCATGGGGAACATAGCAGTTGTCAGAACTTGGGCTTACCAACCAACTTCT 554
558 TTATGACCAACATGAAAGAAATTTTCAACCACTCAAGTCTCTCATCTGAAATATT 617
555 GGGTCCCTGGGTGTGTTTGGCAACTTGTCAATGAAAGCTTATCTTGGAGTCAATCC 614
618 GGGTTCATTTCACTGTTTGGAAATCTTGAACCTTGAACCCGGAATGAGGTGATGATCA 677
615 GGGACACCTCTGCTCCAGACCTGATGAAGAGGTGTACTTGGCTTCACTGAGTGTG 674
678 GAGCAGTGTGATCTGCAATCATGATGATGATGATGATGATGATGATGATGATGATGAT 737
675 GACAGCTGATGCTGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 728
738 GCTCTTATTCGTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 791
```

```
RESULT 6
US-10-027-632-134530/c
; Sequence 134530, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; POLYMORPHISMS IN THE HUMAN GENOME
; FILE REFERENCE: 108827, 129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
```

PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134530
LENGTH: 578
TYPE: DNA
ORGANISM: Human
US-10-027-632-134530

Query Match
Best Local Similarity 6.5%; Score 125.2; DB 15; Length 578;
Matches 139; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 488 GCAGGTTTATGACCCATGCGGAGCAATGAAGCAGTTGGAGGCTACACACAC 547
Db 183 GCAGGCTTTATGACTCCGCGAGAACACAGCAGTTGGAGAACTGACCTACACACAC 124
Qy 548 ACATTCTGGGTCCTCGGCTGGTGGTTTGGCCAACTTGTCAATGAAGGCTTATCTTGAGGT 607
Db 123 ATGTTCTGGGTGCTCGGCTGGTGGTTTGGCCAACTTGTCAATGAAGGCTTGAGGT 64
Qy 608 CCAATCCGGAGACACCTCTCTCCAGACCTGATGAAGAG 649
Db 63 CCAATCCGGAGACCTTATCTCTCCAGACCTGCTGAAGCTG 22

RESULT 7
US-10-029-386-7583/c
Sequence 7583, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AEMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
NUMBER OF SEQ ID NOS: 34286
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 7583
LENGTH: 539
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR19.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUE 9.00e-56
OTHER INFORMATION: NT HIT: g115304094, EVALUE 1.00e-130
OTHER INFORMATION: SWISSPROT HIT: 076090, EVALUE 2.00e-25
US-10-029-386-7583

Query Match
Best Local Similarity 6.4%; Score 122.2; DB 14; Length 539;
Matches 169; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
Qy 247 CTTGCTTGGGTTTCTATGTTACATTTGGTGAAGCCCTGGTGAGGACGAGAA 306

Db 372 CTCCTCGAGGCTTTATGTCAGCTGGTGTGAACCCCTGGTGAAGCCAGTACTATG 313
Qy 307 CTTGCGGTGGCCCGACCGCTCATGATCCAGGTGTAGTTGTGAGAGGCAAGATGA 366
Db 312 CATGCCGTCGCGAGCGCTCATGTGCTGTGGGGGACCGTGCACGACGACGA 253
Qy 367 GGAAGCCGTTTGTCTCGGCGACGCTCATCCGCTACGCGATCTGGGCGCAAGTGTCTAT 426
Db 252 CCGCGCCCGCTTATACCGGCGACACTCATGCGCTACGAGGCGCTCTCGGCGCTCAT 193
Qy 427 CTTGCGGACATCAGGACCTCGGTCTTACAGCGCTTTCCACTTTCACCACTGTGTCT 486
Db 192 CCGCGCTCGTTCAGGACCGCGGTGTTCAGGCGCTTCCCGACCATAGACCACTGTGTGA 133
Qy 487 AGCAGGT 493
Db 132 GGCTGGT 126

RESULT 8
US-10-029-386-21283/c
Sequence 21283, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: AEMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
NUMBER OF SEQ ID NOS: 34286
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21283
LENGTH: 235
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR19.3
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
OTHER INFORMATION: EST HUMAN HIT: AA573517.1, EVALUE 4.00e-56
OTHER INFORMATION: NT HIT: g18923136, EVALUE 1.00e-129
OTHER INFORMATION: SWISSPROT HIT: 076090, EVALUE 1.00e-24
US-10-029-386-21283

Query Match
Best Local Similarity 6.3%; Score 119.8; DB 14; Length 235;
Matches 163; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
Qy 258 GTTTCATGTATCATTTGATGATGAGCGGCTGGTGAAGCCAGTACGAACTTGGCGTGC 317
Db 235 GCTTTTATGTAGCGCTGGTGTGAACCGCTGTGAGACCACTATGATGCGCTGC 176
Qy 318 CCGACCGCTCATGATCCAGGTGTCTAGCTTGTGAGGCGCAAGATGAAGAGCCGTT 377
Db 175 CCGAGCGGCTATGTCGCGGTGGGCGGACCGGTGACGAGCGGCGGCGCGCC 116
Qy 378 TGTGCGGCGGACGCTCATCGCTTACGCGTATCTGGGCGCAAGTGTCTCATCTGCGGACGA 437
Db 115 TCTACCGGCGGACACTCATGCGCTTACGAGGCGCTTCCGCGGTGTCTCATCTGCGCTCG 56
Qy 438 TCAGCAGCTCGGTCTACAAAGCGCTTCCACTTTCACCACTGAGTGTGAGAGG 492
Db 55 TTAGCAGCGGCTGTTCAGGCGCTTCCCGACATAGACCACTGTGTGAGGCTGG 1

RESULT 9
US-10-198-846-11346
Sequence 11346, Application US/10198846

```
Publication No. US20030099974A1
GENERAL INFORMATION:
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIORITY FILING DATE: 2002-07-18
PRIORITY FILING DATE: 2001-07-18
PRIORITY FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11346
LENGTH: 853
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 23, 27, 36, 37, 40, 41,
LOCATION: 43, 50, 51, 54, 57, 853
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-11346
```

```
Query Match
Best Local Similarity 6.0%; Score 114.4; DB 14; Length 853;
Matches 160; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
```

```
QY 256 GGGTTCTATGTTACATGTTGGTAGCCGCTGGTAGAGCACTAGAGAACTTGGCCG 315
DB 272 GGGGTTTATGTTACTGTTAGTGAACCGATGTTGTAATTGGCCCTG 331
QY 316 GCCGACCGCTCATGATCCAGGTGTTAGCTTCTGAGAGGCAAGATGAGAGAGCCG 375
DB 332 GCCAGACAGCTATATGTTCTCTACTCTAGCAATGTTCAAGAGAGAGAGAGAGAG 391
QY 376 TTTGCTGGGCGGACGCTATCCGCTACCGCATCTCGGCGCAAGTGTCTACTCTGGCG 435
DB 392 CCGCTTAGAAGAGACCGCTGATGCGCTACGCTACCTCCCTGCTCATCTTGGCTC 451
QY 436 CATCAGCACTGCTGTACAGGCTTCCCACTTCCACCACTGGTGTAGCAG 491
DB 452 GGTGACACTGCTGTGTGACAAAGATTTCACCAATGAGACCACTGTTGAGAGAG 507
```

```
RESULT 10
US-10-027-632-134529/C
Sequence 134529, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
PRIORITY FILING DATE: 2002-04-30
PRIORITY FILING DATE: 2000-07-12
PRIORITY FILING DATE: 2000-07-12
PRIORITY FILING DATE: 2000-04-20
PRIORITY FILING DATE: 2000-03-29
PRIORITY FILING DATE: 2000-03-29
PRIORITY FILING DATE: 2000-02-24
PRIORITY FILING DATE: 1999-11-23
PRIORITY FILING DATE: 1999-09-28
PRIORITY FILING DATE: 1999-08-09
```

```
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 134529
LENGTH: 748
TYPE: DNA
ORGANISM: Human
US-10-027-632-134529
```

```
Query Match
Best Local Similarity 5.9%; Score 113.2; DB 15; Length 748;
Matches 138; Conservative 0; Mismatches 23; Indels 1; Gaps 1;
```

```
QY 488 GCAGTTTATGATCCATGAGGAGCAATAGCAAGTTGACAGAGTGGGCTTACACACAC 547
DB 688 GCAAGCTTATATGATCTCCGACAGACAGAG-TCGAAACTAGGCTTACACACACAC 630
QY 548 AATTTCTGGTGGTCCCTGGTGTGTTGGCAACTGTCTCATATGAGGCTTATCTTGAGGT 607
DB 629 ATGTTCTGGTGGTCCCTGGTGTGTTGGCAACTGTCTCATATGAGGCTTATCTTGAGGT 570
QY 608 CGAATCCGGAGACCGTCTCTCCAGAGCCGATGATGAG 649
DB 569 CGAATCCGGAGACCGTCTCTCTCCAGAGCCGATGAGGTG 528
```

```
RESULT 11
US-10-198-846-2561/C
Sequence 2561, Application US/10198846
Publication No. US20030099974A1
GENERAL INFORMATION:
```

```
APPLICANT: Lillie, James
APPLICANT: Xu, Yongyao
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
THERAPY OF BREAST CANCER
FILE REFERENCE: MRI-049
CURRENT APPLICATION NUMBER: US/10/198,846
PRIORITY FILING DATE: 2002-07-18
PRIORITY FILING DATE: 2001-07-18
PRIORITY FILING DATE: 2001-07-18
NUMBER OF SEQ ID NOS: 14084
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2561
LENGTH: 462
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 390_462
OTHER INFORMATION: n = A,T,C or G
US-10-198-846-2561
```

```
Query Match
Best Local Similarity 4.9%; Score 94.4; DB 14; Length 462;
Matches 134; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

```
QY 255 TGGGTTTATGTTACTGTTAGTGAACCGATGTTGGAATTTGGCCCT 314
DB 255 TGGGTTTATGTTACTGTTAGTGAACCGATGTTGGAATTTGGCCCT 196
QY 315 GCCCGACCGCTCATGATCCAGGTGTTAGTGTGAGGCAAGATGAGAGAGCC 374
DB 195 GGCAGACAGGCTAATGTTCTCATCTAGAGTGTTCAGAGAGAGAGAGAGAGCC 136
QY 375 GTTGTGGGCGGACGCTCATCCGCTACCGCATCTGGGCAAGTGTCTATCTTGGCA 434
DB 135 GCTGTGTTAGAGAGAGAGGCTGATGCGCTAGCTAATCTACCTCCGCTCATCTTGGCT 76
QY 435 GCATCAGACCTCGGCTTAC 454
DB 75 CGGTGACACTGCTGTGTAC 56
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 01:53:34 ; Search time 8577 Seconds

(without alignments)
1198.340 Million call updates/sec

Title: US-09-622-964-1_COPY_13900_16115

Portect score: 2216
Sequence: 1 attctcctgtcttcttcacg.....ctaatgagtttaataac 2216

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_ptg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_stc: *
28: em_un: *
29: em_vl: *
30: em_hcg_hum: *
31: em_hcg_inv: *
32: em_hcg_other: *
33: em_hcg_mus: *
34: em_hcg_pln: *
35: em_hcg_rtd: *
36: em_hcg_man: *
37: em_hcg_vrc: *
38: em_sy: *
39: em_hcgo_hum: *
40: em_hcgo_mus: *
41: em_hcgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2215	100.0	16125	6	BD136719
2	2188.6	98.8	142082	9	AF139813
3	2188.6	98.8	196080	9	AC004228
4	2187	98.7	112309	9	AC003025
5	2187	98.7	163024	9	AP006260
6	2187	98.7	163915	2	AC084451
7	2187	98.7	163867	2	AP003733
8	2185.4	98.6	133683	2	AC084857
9	1608.4	72.6	160169	2	AC0051664
10	1561.6	70.5	106648	2	AC004588
11	1339.8	60.5	2435	9	BC041664
12	688.4	31.1	706	9	HSWMD2P10
13	662	29.9	726	9	AY357926S10
14	643.8	29.1	1263	6	BD132659
15	643.8	29.1	2210	9	AF057169
16	643.8	29.1	2420	9	AF057170
17	642.2	29.0	1758	6	AX745964
18	642.2	29.0	2171	9	AF073501
19	642.2	29.0	2229	6	BD136720
20	642.2	29.0	2429	6	BD136721
21	642.2	29.0	2441	9	BC015220
22	600.6	27.1	2187	9	AY357925
23	399.4	18.0	424	6	BD059099
24	389.4	17.6	482	9	HSWMD2P11
25	318.6	14.4	1289	4	AY064707
26	318	14.4	356	6	AX908879
27	318	14.4	356	6	BD044412
28	302	13.6	430	9	AY357926S11
29	164	7.4	448	11	G05922
30	164	7.4	1119	9	HUMFERRH
31	164	7.4	1198	6	AX409527
32	164	7.4	1198	6	HUMFERRITH
33	136	6.1	195646	9	AC093709
34	135	6.1	186896	9	AC019057
35	134.4	6.0	181049	9	AC093131
36	133.2	6.0	65042	6	AR235846
37	133.2	6.0	65042	6	AX702088
38	133.2	6.0	149480	6	AX329775
39	133.2	6.0	149480	6	AX329776
40	133.2	6.0	149480	6	AX336193
41	133.2	6.0	149480	9	HU095740
42	133.2	6.0	149480	9	AC130455
43	133.2	6.0	156212	9	HS12684
44	133.2	6.0	175839	9	AC004491
45	133.2	6.0	180399	2	AC079252

ALIGNMENTS

RESULT 1
BD136719 16125 bp DNA linear PAT 18-SEP-2002
LOCUS Best's macular dystrophy gene.
ACCESSION BD136719
VERSION BD136719.1 GI:23231664
KEYWORDS JP 2002504559-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Petrukhin, K., Caskey, T.C., Metzker, M. and Wadelius, C.
TITLE Best's macular dystrophy gene
JOURNAL Patent: JP 2002504559-A 1 (12-FEB-2002)

Pred. No. is the number of results predicted by chance to have a

COMMENT
MERCK & CO INC, CLAES WADELIUS
OS Homo sapiens (human)
PN JP 2002504559-A/1
PD 12-FEB-2002
PF 22-FEB-1999 JP 2000533447
PR 25-FEB-1998 US 60/075941, 18-DEC-1998 US 60/112926 PI
KONSTANTIN PETRUKHIN, THOMAS C CASEKEY, MICHAEL METZKER, CLAES PI
WADELIUS
PC C07K16/18, C07K14/47, C12N5/10, C12N15/09, C12P19/34, C12Q1/66// PC
C12P21/08, C12P21/09, C12N5/00, C12N15/00
CC Best s macular dystrophy gene
FH Key Location/Qualifiers
FT source 1..16125
/organism='Homo sapiens (human)'.
Location/Qualifiers
1..16125
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

ORIGIN
Query Match 100.0%; Score 2215; DB 6; Length 16125;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2216; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCTCTGTTCTTTCCAGCCTGAAACAAAGAGAGATGAGTTCCAGCCCAATCGAGG 60
DB 13900 ATCTCCGTGTTCTTTCCAGCCTGAAACAAAGAGAGATGAGTTCCAGCCCAATCGAGG 13959
QY 61 ACAGAGAGATGCTCAAGCTGGGATCATTTGGCCGCTTCTTAGGCTCTGATGCCATGATC 120
DB 13960 ACAGAGAGATGCTCAAGCTGGGATCATTTGGCCGCTTCTTAGGCTCTGATGCCATGATC 14019
QY 121 ACCATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTTGGGCCCAAGAGGGAATCCCTTC 180
DB 14020 ACCATCTCTCCAGAGGCAAACTCAAGAGCAAACTACTGTTGGGCCCAAGAGGGAATCCCTTC 14079
QY 181 TCACAGAGGCGCTGCCAAAACACAGAGCGAGCCAAACAGACGTTAGGGCCAGAGAG 240
DB 14080 TCACAGAGGCGCTGCCAAAACACAGAGCGAGCCAAACAGACGTTAGGGCCAGAGAG 14139
QY 241 ACAACAGAGCGCTGAAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCCCTGTATCA 300
DB 14140 ACAACAGAGCGCTGAAAGCTTAAAGCTGTGAGCGCTTCAAGTCTGCCCCCTGTATCA 14199
QY 301 GGCAGGCTACTACAGTSCCCCAAGAGCGCCCTCAAGCCCACTCCCATGTTCTCCCC 360
DB 14200 GGCAGGCTACTACAGTSCCCCAAGAGCGCCCTCAAGCCCACTCCCATGTTCTCCCC 14259
QY 361 TAGAACATCAAGCGCCGCTCAAGAGCTTCAAGTGTCAAGGCAATGACCAAGACAAA 420
DB 14260 TAGAACATCAAGCGCCGCTCAAGAGCTTCAAGTGTCAAGGCAATGACCAAGACAAA 14319
QY 421 GCTTAAAGACTGTAGTCTTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGACGATG 480
DB 14320 GCTTAAAGACTGTAGTCTTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGACGATG 14379
QY 481 GGGCTTGTATGAGCAACCAAGAGTATCTCAAGTGTGAGGAGAAACTGTGAGTTAAAC 540
DB 14380 GGGCTTGTATGAGCAACCAAGAGTATCTCAAGTGTGAGGAGAAACTGTGAGTTAAAC 14439
QY 541 TGACGGATGTCGAGAGATCCCGGAATATCATTAAGAACTTTTGGAAATCAACCA 600
DB 14440 TGACGGATGTCGAGAGATCCCGGAATATCATTAAGAACTTTTGGAAATCAACCA 14499
QY 601 CCAACATACACTACACTCAAGATCAATGGAATCTTATTTGGCTTGGAAAAAGGT 660
DB 14500 CCAACATACACTACACTCAAGATCAATGGAATCTTATTTGGCTTGGAAAAAGGT 14559
QY 661 CTGTCTCACTGAAACCAAGGGGCACTGCAATTCCTGCTGCCCCCAGCTTCCCTTG 720
DB 14560 CTGTCTCACTGAAACCAAGGGGCACTGCAATTCCTGCTGCCCCCAGCTTCCCTTG 14619

QY 721 CTGTGAGCTTACCTTCTCCACAAATTTCTTAGAGTTTCCATCTGCGCAGAGCACTGG 780
DB 14620 CTGTGAGCTTACCTTCTCCACAAATTTCTTAGAGTTTCCATCTGCGCAGAGCACTGG 14679
QY 781 ACCATCCGCCAGCACTGGCTTGGGGTATATATCTTGGCACTTTCACAGGAGATCTTAGGGA 840
DB 14680 ACCATCCGCCAGCACTGGCTTGGGGTATATATCTTGGCACTTTCACAGGAGATCTTAGGGA 14739
QY 841 AGTGTGGGAGCTTTTCTACCTTCACTCCGATATACCCGGAAGACTTCTTGGAGCAG 900
DB 14740 AGTGTGGGAGCTTTTCTACCTTCACTCCGATATACCCGGAAGACTTCTTGGAGCAG 14799
QY 901 GTGAAGAAATAGATGTTGTCTGACAGAAATGCTGTGAGAGAACTGCCCAAGGCTGAC 960
DB 14800 GTGAAGAAATAGATGTTGTCTGACAGAAATGCTGTGAGAGAACTGCCCAAGGCTGAC 14859
QY 961 AGGCGAGCTTAGCTAGAGAGATGTTATCACTGGGCCCACTTACTTTAGCAAGGCTG 1020
DB 14860 AGGCGAGCTTAGCTAGAGAGATGTTATCACTGGGCCCACTTACTTTAGCAAGGCTG 14919
QY 1021 CTGACCCAAACCATAGAGGTGGCACTGAGTGAATGACAGATGAACATTTCCCATAC 1080
DB 14920 CTGACCCAAACCATAGAGGTGGCACTGAGTGAATGACAGATGAACATTTCCCATAC 14979
QY 1081 TATTTAGGTTAGTACCCAAAGCACTACAGGAAAGGCTGGCAAGAACTGCTCACTAGG 1140
DB 14980 TATTTAGGTTAGTACCCAAAGCACTACAGGAAAGGCTGGCAAGAACTGCTCACTAGG 15039
QY 1141 AACTGTGATAGTGTGAGGTTGAGGGTTCAGCGCCCTTAGTGTATTTCTCATGCTCG 1200
DB 15040 AACTGTGATAGTGTGAGGTTGAGGGTTCAGCGCCCTTAGTGTATTTCTCATGCTCG 15099
QY 1201 GGAACCTCACCAAAATCTTCTGCTTCTTGGGCTCAGCCCAAGCTGTCAAAATCA 1260
DB 15100 GGAACCTCACCAAAATCTTCTGCTTCTTGGGCTCAGCCCAAGCTGTCAAAATCA 15159
QY 1261 GATATTTCCCTTATTTCCAGATTTCTGCACTGACACCAATTTAAACCCCACTT 1320
DB 15160 GATATTTCCCTTATTTCCAGATTTCTGCACTGACACCAATTTAAACCCCACTT 15219
QY 1321 CAGCCCAATCAAGTGGAGAGAGTAACTTCTCTTCTTGTGATTTCTCAAGCACTACT 1380
DB 15220 CAGCCCAATCAAGTGGAGAGAGTAACTTCTCTCTTGTGATTTCTCAAGCACTACT 15279
QY 1381 TTCAAGGCTCAAGACGAGCTATTATGATGAACCTTAAAGGGAACAATTTGANT 1440
DB 15280 TTCAAGGCTCAAGACGAGCTATTATGATGAACCTTAAAGGGAACAATTTGANT 15339
QY 1441 CTGCTTCTTAGGCTAAGACAGAACTTGGCAAACTGTGGCTGTTCAGCAAGATG 1500
DB 15340 CTGCTTCTTAGGCTAAGACAGAACTTGGCAAACTGTGGCTGTTCAGCAAGATG 15399
QY 1501 TTCAATTTTAAAGATCTTGTCTTGGGCTGGGTGTGAGGCAAGGATTCACAGAGGTCA 1560
DB 15400 TTCAATTTTAAAGATCTTGTCTTGGGCTGGGTGTGAGGCAAGGATTCACAGAGGTCA 15459
QY 1561 GGAAGTTGAGCAACCTGGGCCAAGATGATGAACCCCATCTTCAACCAAAAAATACAA 1620
DB 15460 GGAAGTTGAGCAACCTGGGCCAAGATGATGAACCCCATCTTCAACCAAAAAATACAA 15519
QY 1621 ATCACTGGCGGTGTGTGTGCTGTAGTCCCAAGCAAGAGATTTGAGGGAGAAATTGC 1680
DB 15520 ATCACTGGCGGTGTGTGTGCTGTAGTCCCAAGCAAGAGATTTGAGGGAGAAATTGC 15579
QY 1681 TTGAACCAAGAGAGTGTGTGTGAGATTTGAGCACTGCAATTCAGCTGGCGAGC 1740
DB 15580 TTGAACCAAGAGAGTGTGTGTGAGATTTGAGCACTGCAATTCAGCTGGCGAGC 15639
QY 1741 GGAAGTGAAGTGTCTCAAAAAAAGAGATGCTCAACTTTGGCTTCTACT 1800
DB 15640 GGAAGTGAAGTGTCTCAAAAAAAGAGATGCTCAACTTTGGCTTCTACT 15699

Qy	1801	GCAACATTTTGGATTGTAATGAAAGAGTACCTTCATCTACTTATGCTGTTAATCTTTCAT	1860
Db	15700	GCAACATTTTGGATTGTAATGAAAGTACCTTCATCTACTTATGCTGTTAATCTTTCAT	15759
Qy	1861	TTTCACAGGAGGATGAGACACATTCCTTAACCTGCTTCTTAATGGGAGTCTTCCGACCCA	1920
Db	15760	TTTCACAGGAGGATGAGACACATTCCTTAACCTGCTTCTTAATGGGAGTCTTCCGACCCA	15819
Qy	1921	GGTCTCCTACCTGATGTGTGACACGACGAGACATGATCAGTCAAGCCATTCAGCTGTCC	1980
Db	15820	GGTCTCCTACCTGATGTGTGACACGACGAGACATGATCAGTCAAGCCATTCAGCTGTCC	15879
Qy	1981	ACACTGAAAGACGTGTCTCTCAACAGCCTGATCAATCAATGGTTAGCTTAATGATTAATAAT	2040
Db	15880	ACACTGAAAGACGTGTCTCTCAACAGCCTGATCAATCAATGGTTAGCTTAATGATTAATAAT	15939
Qy	2041	CCCACTACTTCAAGCCTTTAATGCTTTTATTCATTAATAAAGCTGTGAAGCTAGACTGAA	2100
Db	15940	CCCACTACTTCAAGCCTTTAATGCTTTTATTCATTAATAAAGCTGTGAAGCTAGACTGAA	15999
Qy	2101	CCATTGGAAACATTTAATCTGAGACTGTGATTCAGAGTCGGGAAACCTTGATCTAATCTG	2160
Db	16000	CCATTGGAAACATTTAATCTGAGACTGTGATTCAGAGTCGGGAAACCTTGATCTAATCTG	16059
Qy	2161	AATCAAGACAGCCACACTTAACTGATATCTGCCCAACTAATGATTAATAATATAC	2216
Db	16060	AATCAAGACAGCCACACTTAACTGATATCTGCCCAACTAATGATTAATAATATAC	16115
RESULT 2			
AF139813/c			
LOCUS	142092 bp	DNA	linear PRI 22-APR-1999
DEFINITION	Homo sapiens clone pDJ759j12 chromosome 11 map 11q13, complete sequence.		
ACCESSION	AF139813		
VERSION	AF139813.1	GI:4633837	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE	1 (bases 1 to 142092)		
AUTHORS	Petrushkin,K., Koishi,M.J., Bakall,B., Li,W., Xie,G., Marknell,T., Sandgren,O., Forsman,K., Holmgren,G., Andreasson,S., Vojic,M., Metzger,M.U., Caskey,C.T. and Wadelius,C.		
	Identification of the gene responsible for Best macular dystrophy Nat. Genet. 19 (3), 241-247 (1998)		
TITLE	98324772		
JOURNAL	MEDLINE		
PUBMED	9662395		
REFERENCE	2 (bases 1 to 142092)		
AUTHORS	McGarty-Dugan,V.A., Hammond,H.A., Clement,M.K., Larson,D.R., Liu,X., Soderman,A.R., McGowan,J.M., DeAngelis,D.M., Ian,C., Fitzpatrick,E.S., Harrison,K.M., Petrushkin,K., Caskey,C.T. and Metzger,M.U.		
	Direct Submission		
TITLE	Submitted (01-APR-1999) Department of Human Genetics, Merck & Co., Inc., Summerville Pike, West Point, PA 19486, USA		
JOURNAL	IMPORTANT: This submission contains the entire insert of clone pDJ759j12 which comes from a PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. This sequence has been finished such that all consensus base calls consist of two or more separate clones with double-stranded coverage or two or more separate clones with two types of sequencing chemistry. The expected phred/blah calculated errors/10kb is 0.00.		
FEATURES	location/Qualifiers		
source	1. 142092		
	/organism="Homo sapiens"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9606"		
	/chromosome="11"		
	/map="11q13"		
	/clone="pDJ759j12"		
	complement (19..1855)		
	/rpt_family="L1"		
	repeat_region		
	1861..2001		
	/rpt_family="L1"		
	repeat_region		
	1990..2317		
	/rpt_family="Alu"		
	complement (2031..2257)		
	/rpt_family="SVA"		
	repeat_region		
	2320..2974		
	/rpt_family="L1"		
	repeat_region		
	2985..3540		
	/rpt_family="Alu"		
	complement (4034..4124)		
	/rpt_family="MIR"		
	repeat_region		
	4417..4995		
	/rpt_family="Alu"		
	complement (4432..4936)		
	/rpt_family="SVA"		
	repeat_region		
	5569..6095		
	/rpt_family="SVA"		
	complement (7239..7542)		
	/rpt_family="Alu"		
	repeat_region		
	7306..7497		
	/rpt_family="SVA"		
	repeat_region		
	7752..7828		
	/rpt_family="MIR"		
	repeat_region		
	9099..9403		
	/rpt_family="Alu"		
	complement (9526..9897)		
	/rpt_family="MER1"		
	repeat_region		
	9635..9752		
	/rpt_family="MER1B"		
	repeat_region		
	11924..12197		
	/rpt_family="Alu"		
	complement (11955..12151)		
	/rpt_family="SVA"		
	repeat_region		
	13466		
	/rpt_family="Alu"		
	repeat_region		
	16757..17039		
	/rpt_family="Alu"		
	complement (16775..17010)		
	/rpt_family="SVA"		
	repeat_region		
	18220		
	/rpt_family="Alu"		
	complement (18220..18507)		
	repeat_region		
	18225..18485		
	/rpt_family="SVA"		
	gene		
	20101..23446		
	/gene="HUMFERRIT1"		
	/note="HUMAN ferritin heavy chain gene"		
	join(20101..20421,22218..22364,22621..22746,22842..23446)		
	/gene="HUMFERRIT1"		
	/product="ferritin heavy chain"		
	repeat_region		
	21506..21714		
	/rpt_family="Alu"		
	repeat_region		
	21506..21714		
	/rpt_family="SVA"		
	complement (23272..37375)		
	/gene="VMD2"		
	/note="Best macular dystrophy gene"		
	28155..28255,29427..29579,30274..30351,30740..30894,		
	31787..32020,32537..32631,35777..35964,37308..37375)		
	/gene="VMD2"		
	/product="Best macular dystrophy protein"		
	complement (23735..23954)		
	/rpt_family="Alu"		
	repeat_region		
	23789..23956		
	/rpt_family="SVA"		

```
repeat_region 26429..26707
/rpt_family="Alu"
complement(26444..26696)
repeat_region /rpt_family="SVA"
27103..27349
/rpt_family="Alu"
complement(27106..27323)
repeat_region /rpt_family="SVA"
complement(29161..29307)
repeat_region /rpt_family="Alu"
complement(29762..30043)
repeat_region /rpt_family="Alu"
29786..30017
/rpt_family="SVA"
complement(30462..30516)
repeat_region /rpt_family="MIR"
complement(31095..31409)
repeat_region /rpt_family="Alu"
33005..33260
/rpt_family="Alu"
complement(33247..33597)
repeat_region /rpt_family="LMC2"
complement(33849..34431)
repeat_region /rpt_family="Alu"
34076..34411
/rpt_family="SVA"
complement(34640..34890)
repeat_region /rpt_family="Alu"
34972..35401
complement(34988..35392)
repeat_region /rpt_family="SVA"
36236..36487
/rpt_family="Alu"
complement(36253..36457)
repeat_region /rpt_family="SVA"
complement(38661..38855)
repeat_region /rpt_family="L1"
complement(40997..41224)
repeat_region /rpt_family="Alu"
41056..41213
/rpt_family="SVA"
42745..43020
/rpt_family="Alu"
complement(42763..42966)
repeat_region /rpt_family="SVA"
43312..43457
complement(43963..44217)
repeat_region /rpt_family="L1"
complement(45296..45533)
repeat_region /rpt_family="SVA"
complement(45300..45401)
repeat_region /rpt_family="LTR5"
complement(45407..45533)
repeat_region /rpt_family="LTR5"
complement(45542..45818)
repeat_region /rpt_family="Alu"
45598..45770
/rpt_family="SVA"
complement(45862..46004)
repeat_region /rpt_family="L1"
complement(45998..46282)
repeat_region /rpt_family="Alu"
46022..46260
/rpt_family="SVA"
complement(46692..47583)
repeat_region /rpt_family="Alu"
46885..47215
/rpt_family="SVA"
47363..47571
complement(47665..47773)
repeat_region /rpt_family="SVA"
```

```
repeat_region /rpt_family="Alu"
complement(47792..48077)
repeat_region /rpt_family="Alu"
complement(48110..48388)
repeat_region /rpt_family="L1"
complement(48827..49102)
repeat_region /rpt_family="Alu"
48883..49110
/rpt_family="SVA"
complement(49395..49626)
repeat_region /rpt_family="Alu"
50184..50437
```

Query Match 98.8%; Score 2188.6; DB 9; Length 142092;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2205; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
QY 1 ATCTCCTGTTCTTTTCAGCCTGAACAAAGAGATGAGTTCCAGGCCAATCAGAG 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25497 ATCTCCTGTTCTTTTCAGCCTGAACAAAGAGATGAGTTCCAGGCCAATCAGAG 25438

QY 61 ACAGAGAGATGCTCAACGCTGACATCTTGGCCCTTCTTAGCCTGATCCATGATC 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25437 ACAGAGAGATGCTCAACGCTGACATCTTGGCCCTTCTTAGCCTGATCCATGATC 25378

QY 121 ACCATCTTCCAGGGCAACTCAAGACCAACTAGTGTGCCCCAAGAGGGAATCCCTTC 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25377 ACCATCTTCCAGGGCAACTCAAGACCAACTAGTGTGCCCCAAGAGGGAATCCCTTC 25318

QY 181 TCCACAGAGGCGCTGCCCAAAACCAAGAGGCCCAACAGACGTTAGGCGCCAGAG 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25317 TCCACAGAGGCGCTGCCCAAAACCAAGAGGCCCAACAGACGTTAGGCGCCAGAG 25258

QY 241 ACAACAAGCCTGGAAGCTTAGGCTGTGACGCTTCAAGTCTGCCCACTGATACGA 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25257 ACAACAAGCCTGGAAGCTTAGGCTGTGACGCTTCAAGTCTGCCCACTGATACGA 25198

QY 301 GGGCGAGGTACTAGAGGCCCAAGACGCCCTCAGCCTTCAAGTCTGCCCACTGATACGA 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25197 GGGCGAGGTACTAGAGGCCCAAGACGCCCTCAGCCTTCAAGTCTGCCCACTGATACGA 25138

QY 361 TAGAACATCAGCGCCCTCAAGCTTCAAGTGTCAAGGTGACAGGACACCAAGACAA 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25137 TAGAACATCAGCGCCCTCAAGCTTCAAGTGTCAAGGTGACAGGACACCAAGACAA 25078

QY 421 GCTTAAGACCTGAGGCTTCTGGGCGCAAGAAAGTTTGAATTGCTTCAGAGAGCATG 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25077 GCTTAAGACCTGAGGCTTCTGGGCGCAAGAAAGTTTGAATTGCTTCAGAGAGCATG 25018

QY 481 GGGCTTGTATGAGACCCAGCAAGTATCTCAAGTGTGAGGAAAGCTGTGAGTTAAC 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 25017 GGGCTTGTATGAGACCCAGCAAGTATCTCAAGTGTGAGGAAAGCTGTGAGTTAAC 24958

QY 541 TGAAGATATGCGAGAGATCCCGGAAATCAGCTCAAGAAAGCTTTGGAACATACCA 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24957 TGAAGATATGCGAGAGATCCCGGAAATCAGCTCAAGAAAGCTTTGGAACATACCA 24898

QY 601 CCAACATACACTACCTCAAGATCAATGATCCTTATTTGGCCCTTGAAGAAAGCT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24897 CCAACATACACTACCTCAAGATCAATGATCCTTATTTGGCCCTTGAAGAAAGCT 24838

QY 661 CAGTCTCTCACTGAACAGGGGCACTGATTTGCCCTTGGCCCAAGCTTCCCTTG 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24837 CAGTCTCTCACTGAACAGGGGCACTGATTTGCCCTTGGCCCAAGCTTCCCTTG 24778

QY 721 CTTGAGGCTACCTTCTTCAATTTCCAGAGGTTCCATCATTTGCCAAGACACTGG 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24777 CTTGAGGCTACCTTCTTCAATTTCCAGAGGTTCCATCATTTGCCAAGACACTGG 24718

QY 781 ACCTACGCCACGACTGGCTTGGGGTATATCTTGGCACTTCAAGGATCTTAGGGA 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 24717 ACCTACGCCACGACTGGCTTGGGGTATATCTTGGCACTTCAAGGATCTTAGGGA 24658
```

QY 841 AGTGTGGGACCTTTTCTCACTTCACTGATATCACTGGGAGAGACTTCTTGGGACAG 900
 Db 24657 AGTGTGGGACCTTTTCTCACTTCACTGATATCACTGGGAGAGACTTCTTGGGACAG 24598
 QY 901 GTGAGGAGATGAGTGTGTGCTGACCAAGATGCTGTGAGAACTGCCCCAGGGCTGAC 960
 Db 24597 GTGAGGAGATGAGTGTGTGCTGACCAAGATGCTGTGAGAACTGCCCCAGGGCTGAC 24538
 QY 961 AGGCGAGGCTTAGCTGAGAGATGTTTCACTGGGCCCACTTACTTGAAGCAAGGGTGG 1020
 Db 24537 AGGCGAGGCTTAGCTGAGAGATGTTTCACTGGGCCCACTTACTTGAAGCAAGGGTGG 24478
 QY 1021 CTGACCCAAAACATGAGTGTGAGTCACTGATGATGACATGACACTTCCCCCATAC 1080
 Db 24477 CTGACCCAAAACATGAGTGTGAGTCACTGATGATGACATGACACTTCCCCCATAC 24418
 QY 1081 TATTTAGGGTAGTACCCAGACATAAGAAAGGGTGGAGAACTGCTCACTCTAGG 1140
 Db 24417 TATTTAGGGTAGTACCCAGACATAAGAAAGGGTGGAGAACTGCTCACTCTAGG 24358
 QY 1141 AACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 Db 24357 AACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 24298
 QY 1201 GGAACCTCACCAAAATCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
 Db 24297 GGAACCTCACCAAAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 24238
 QY 1261 GATATTTCCCTTTATTCAGATTTCTGACACTGTCACCAATATATAAACACCCCACTT 1320
 Db 24237 GATATTTCCCTTTATTCAGATTTCTGACACTGTCACCAATATATAAACACCCCACTT 24178
 QY 1321 CACCCCCAATCAGTGGGAGAGAGTGTATCTCCCTTTTGTGATTTCTCAAGAGTACT 1380
 Db 24177 CACCCCCAATCAGTGGGAGAGAGTGTATCTCCCTTTTGTGATTTCTCAAGAGTACT 24118
 QY 1381 TTGACGGGTGAGAACGACGACTATTTGATTTGAACTTTAAAGGGCAACAATTTGANT 1440
 Db 24117 TTGACGGGTGAGAACGACGACTATTTGATTTGAACTTTAAAGGGCAACAATTTGANT 24059
 QY 1441 CTTGCTTTAGGCTAGACAGAACTTTGGCAACATCTGTGGCTGTTCAGCAAGAGATG 1500
 Db 24058 CTTGCTTTAGGCTAGACAGAACTTTGGCAACATCTGTGGCTGTTCAGCAAGAGATG 23999
 QY 1501 TTGATATTTAAAGATCTTGTCTTGGCTGTGGCTGTGGAGCAATGATTCACAGAGTCA 1560
 Db 23998 TTGATATTTAAAGATCTTGTCTTGGCTGTGGCTGTGGAGCAATGATTCACAGAGTCA 23939
 QY 1561 GGAGTTTGAGACCACTGGCCAAATGATGAAACCCCATCTCTACCAAAAAAATATCA 1620
 Db 23938 GGAGTTTGAGACCACTGGCCAAATGATGAAACCCCATCTCTACCAAAAAAATATCA 23879
 QY 1621 ATCAGCTGCGCTGT 1680
 Db 23878 ATCAGCTGCGCTGT 23819
 QY 1681 TTGACCCAGAGAGT 1740
 Db 23818 TTGACCCAGAGAGT 23759
 QY 1741 GGAGTGAGACTGTCTCAAAAAAAGATGATGATGATGATGATGATGATGATGATGATG 1800
 Db 23758 GGAGTGAGACTGTCTCAAAAAAAGATGATGATGATGATGATGATGATGATGATGATG 23699
 QY 1801 GCAAACTTTTGTATTTGAAATGAGGTATCTTCAATCTTATGCTGTATATATCTTTGAT 1860
 Db 23698 GCAAACTTTTGTATTTGAAATGAGGTATCTTCAATCTTATGCTGTATATATCTTTGAT 23639
 QY 1861 TCTCACAAGGAGAGAGCAATCTTCACTGCTCTCTATATGGGAGTGTGTGTGTGTGTGTGT 1920
 Db 23638 TCTCACAAGGAGAGAGCAATCTTCACTGCTCTCTATATGGGAGTGTGTGTGTGTGTGTGT 23579
 QY 1921 GGTCTCACAAGT 1980

Db 23578 GGTCTCACAAGT 23519
 QY 1981 ACACTGAAGACCTGTCTTCAACACAGCTGAAATCAATGTTTGTGTTAATGATTAAT 2040
 Db 23518 ACACTGAAGACCTGTCTTCAACACAGCTGAAATCAATGTTTGTGTTAATGATTAAT 23459
 QY 2041 CCCAGACTACTTACAGCTTATATGCTTTTATTCATTAATGATGTAAGTACTGAA 2100
 Db 23458 CCCAGACTACTTACAGCTTATATGCTTTTATTCATTAATGATGTAAGTACTGAA 23399
 QY 2101 CCATTGGAAACATTTAATCTGACACTGTGATTCAGAGTGGGAAACCTTATGTTCTG 2160
 Db 23398 CCATTGGAAACATTTAATCTGACACTGTGATTCAGAGTGGGAAACCTTATGTTCTG 23339
 QY 2161 AATCCAAAGACGACACCTTATATGCTTCCCAATTAATGATTAATTAATAC 2216
 Db 23338 AATCCAAAGACGACACCTTATATGCTTCCCAATTAATGATTAATTAATAC 23283
 RESULT 3
 AC004228 196080 bp DNA linear PRI 24-FEB-1999
 LOCUS
 DEFINITION
 Homo sapiens Chromosome 11q12.2 PAC clone pdu519013 containing human gene for ferritin heavy chain (FTH), complete sequence.
 AC004228
 VERSION
 AC004228.2 GI:4263838
 KEYWORDS
 HTG.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 196080)
 2 (bases 1 to 196080)
 REFERENCES
 Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brigiac,S., Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Loo,H., Loo,K.N., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
 Direct Submission
 Submitted (26-FEB-1998) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235-8591, USA
 TITLE
 JOURNAL
 REFERENCES
 Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M., Buettner,J., Bumester,R., Card,P., deSaillobat,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,J., Schultz,R.A., Stimson,S., Waller,K. and Ward,T.
 Direct Submission
 Submitted (24-FEB-1999) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75235-8591, USA
 COMMENT
 On Feb 24, 1999 this sequence version replaced gi:291173.
 IMPORTANT: This submission contains the entire insert of clone pdu519013. pdu519013 comes from the RPI-3 PAC library constructed at the Roswell Park Cancer Institute by the Pieter de Jong group.
 CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11p12.2 Best's disease region mapped between STS D11S461 and EST AHNAK. This region spans over 1.5 Mbp.
 MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA), STS8 D11S639 and WI-7524
 MAPPED CLONE OVERLAP: PACs pdu466a11 and pdu756b9.

FEATURES		Location/Qualifiers
source	1.	.196080
		/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
repeat_region		complement(2210, .2482)
repeat_region		/rpt_family="Alu"
repeat_region		3457, .3743
repeat_region		/rpt_family="Alu"
repeat_region		complement(20145, .20758)
repeat_region		/rpt_family="Alu"
repeat_region		complement(23194, .23599)
repeat_region		/rpt_family="Alu"
repeat_region		complement(27367, .27685)
repeat_region		/rpt_family="Alu"
repeat_region		28047, .28457
repeat_region		/rpt_family="Alu"
repeat_region		complement(29390, .29565)
repeat_region		/rpt_family="Alu"
repeat_region		complement(31076, .31255)
repeat_region		/rpt_family="Alu"
repeat_region		complement(34948, .35045)
repeat_region		/rpt_family="Alu"
repeat_region		35232, .35538
repeat_region		/rpt_family="Alu"
repeat_region		complement(36242, .36540)
repeat_region		/rpt_family="Alu"
repeat_region		37238, .37522
repeat_region		/rpt_family="Alu"
repeat_region		complement(38555, .38650)
repeat_region		/rpt_family="Alu"
repeat_region		complement(40483, .41164)
repeat_region		/rpt_family="Alu"
repeat_region		42042, .42116
repeat_region		/rpt_family="MIR"
repeat_region		42677, .42817
repeat_region		/rpt_family="L1"
repeat_region		42856, .43157
repeat_region		/rpt_family="L1"
repeat_region		43148, .43380
repeat_region		/rpt_family="L1"
repeat_region		complement(43734, .43999)
repeat_region		/rpt_family="L1"
repeat_region		complement(46998, .47169)
repeat_region		/rpt_family="MIR"
repeat_region		complement(48048, .48346)
repeat_region		/rpt_family="Alu"
repeat_region		48482, .48737
repeat_region		/rpt_family="Alu"
repeat_region		complement(48893, .49037)
repeat_region		/rpt_family="L1"
repeat_region		49118, .49292
repeat_region		/rpt_family="Alu"
repeat_region		49791, .50059
repeat_region		/rpt_family="Alu"
repeat_region		50436, .50501
repeat_region		/rpt_family="Alu"
repeat_region		complement(51147, .51441)
repeat_region		/rpt_family="Alu"
repeat_region		complement(52700, .52979)
repeat_region		/rpt_family="Alu"
repeat_region		53281, .53568
repeat_region		/rpt_family="Alu"
repeat_region		54071, .54204
repeat_region		/rpt_family="Alu"
repeat_region		54506, .54793
repeat_region		/rpt_family="Alu"
repeat_region		54807, .55061
repeat_region		/rpt_family="Alu"
repeat_region		complement(64533, .64826)
repeat_region		/rpt_family="Alu"
repeat_region		complement(65245, .65783)
repeat_region		/rpt_family="Alu"

repeat_region		complement(68465, .68732)
repeat_region		/rpt_family="Alu"
repeat_region		complement(70191, .70476)
repeat_region		/rpt_family="Alu"
repeat_region		70560, .70692
repeat_region		/rpt_family="Alu"
repeat_region		72294, .72470
repeat_region		/rpt_family="Alu"
repeat_region		complement(72800, .73095)
repeat_region		/rpt_family="Alu"
repeat_region		73424, .73695
repeat_region		/rpt_family="Alu"
repeat_region		73784, .74071
repeat_region		/rpt_family="Alu"
repeat_region		74652, .74935
repeat_region		/rpt_family="Alu"
repeat_region		complement(75265, .75514)
repeat_region		/rpt_family="Alu"
repeat_region		complement(76086, .76353)
repeat_region		/rpt_family="Alu"
repeat_region		complement(78038, .78312)
repeat_region		/rpt_family="Alu"
repeat_region		78513, .78734
repeat_region		/rpt_family="Alu"
repeat_region		78866, .79180
repeat_region		/rpt_family="Alu"
repeat_region		complement(79384, .79673)
repeat_region		/rpt_family="Alu"
repeat_region		complement(81398, .81690)
repeat_region		/rpt_family="Alu"
repeat_region		82013, .82288
repeat_region		/rpt_family="Alu"
repeat_region		complement(83419, .83600)
repeat_region		/rpt_family="MIR"
repeat_region		complement(87320, .87624)
repeat_region		/rpt_family="Alu"
repeat_region		complement(87710, .87987)
repeat_region		/rpt_family="Alu"
repeat_region		complement(88186, .88472)
repeat_region		/rpt_family="Alu"
repeat_region		88822, .89049
repeat_region		/rpt_family="Alu"
repeat_region		89224, .89308
repeat_region		/rpt_family="MIR20"
repeat_region		89725, .90487
repeat_region		/rpt_family="Alu"
repeat_region		90504, .90618
repeat_region		/rpt_family="Alu"
repeat_region		complement(93313, .93952)
repeat_region		/rpt_family="Alu"
repeat_region		96938, .97517
repeat_region		/rpt_family="Alu"
repeat_region		complement(104199, .104514)
repeat_region		/rpt_family="Alu"
repeat_region		complement(109738, .109979)
repeat_region		/rpt_family="Alu"
repeat_region		110022, .110141
repeat_region		/rpt_family="Alu"
repeat_region		complement(111091, .112016)
repeat_region		/rpt_family="Alu"
repeat_region		113210, .113419
repeat_region		/rpt_family="L1"
repeat_region		complement(114951, .115005)
repeat_region		/rpt_family="MIR"
repeat_region		115685, .115791
repeat_region		/rpt_family="MIR"
repeat_region		116870, .117005
repeat_region		/rpt_family="MIR"
repeat_region		complement(117877, .118172)
repeat_region		/rpt_family="Alu"
repeat_region		118352, .118600
repeat_region		/rpt_family="Alu"
repeat_region		118679, .118813

/rpe_family="Alu"
complemente(121233..121520)
repeat_region /rpe_family="Alu"
Query Match 98.8%; Score 2188.6; DB 9; Length 196080;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2205; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
QY 1 ATCTCCGTTTCTTTCCAGCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGG 60
DB 191309 ATCTCCGTTTCTTTCCAGCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGG 191368
QY 61 ACAGAGAGATGCTCAAGCTGAGCATCATTTGGCCCTTCTAGGCTCTCATGCCATGATC 120
DB 191369 ACAGAGAGATGCTCAAGCTGAGCATCATTTGGCCCTTCTAGGCTCTCATGCCATGATC 191428
QY 121 ACCATCTCCCAAGGCAAACTCAAGACCAAACTACTGTGAGCCCAAGAGGAAATCCCTTC 180
DB 191429 ACCATCTCCCAAGGCAAACTCAAGACCAAACTACTGTGAGCCCAAGAGGAAATCCCTTC 191488
QY 181 TCCACAGAGGCTGCCCCAAACCAACAGGACGCCAAACAGAAAGTTAAAGGCGCAGAGAG 240
DB 191489 TCCACAGAGGCTGCCCCAAACCAACAGGACGCCAAACAGAAAGTTAAAGGCGCAGAGAG 191548
QY 241 ACAACAAGGCTGGAAGCTTAAGGCTGAGACGCTTCAAGTCTGAGCCCACTGATCAGA 300
DB 191549 ACAACAAGGCTGGAAGCTTAAGGCTGAGACGCTTCAAGTCTGAGCCCACTGATCAGA 191608
QY 301 GGGCAGGCTACTACAGTGCCCAACAGAGCCCTCAGGCCCACTCCCATGTTCTTCCCC 360
DB 191609 GGGCAGGCTACTACAGTGCCCAACAGAGCCCTCAGGCCCACTCCCATGTTCTTCCCC 191668
QY 361 TAGAACCTATAGGCGCGCTCAAAAGTTCTACAGTGTCAAGGCAATAGACACCAAGACAAA 420
DB 191669 TAGAACCTATAGGCGCGCTCAAAAGTTCTACAGTGTCAAGGCAATAGACACCAAGACAAA 191728
QY 421 GCTTAAAGACTGTGAGTGTGGGAGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATG 480
DB 191729 GCTTAAAGACTGTGAGTGTGGGAGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATG 191788
QY 481 GGGCCTTGATGAGCAACCAAGATATCTCAAGTGAAGAGAAAAGTTGAGATTAAAC 540
DB 191789 GGGCCTTGATGAGCAACCAAGATATCTCAAGTGAAGAGAAAAGTTGAGATTAAAC 191848
QY 541 TGAAGGATATGCGAGAGATCCCGGAAATCACTCAAGAAAGCTTTGAAACATTCACCA 600
DB 191849 TGAAGGATATGCGAGAGATCCCGGAAATCACTCAAGAAAGCTTTGAAACATTCACCA 191908
QY 601 CCAACATPACACATACCTCAAGATCATGTGATCTTATTTGGGCTTTGAAACACAGT 660
DB 191909 CCAACATPACACATACCTCAAGATCATGTGATCTTATTTGGGCTTTGAAACACAGT 191968
QY 661 CTGTCTCCACCTGAAACAGAGGCACTGATGCGCTGTGCCCCACCCCACTTCCCTTG 720
DB 191969 CTGTCTCCACCTGAAACAGAGGCACTGATGCGCTGTGCCCCACCCCACTTCCCTTG 192028
QY 721 CTGTGAGCTACCTCTCTCCACAAATTTCTAGGGTTCATCACTGCGCAAGACACTGG 780
DB 192029 CTGTGAGCTACCTCTCTCCACAAATTTCTAGGGTTCATCACTGCGCAAGACACTGG 192088
QY 781 ACCTAGGCCACAGACTGTGGGATATATCTTGGCACTTCAAGAGGATCTTGGGA 840
DB 192089 ACCTAGGCCACAGACTGTGGGATATATCTTGGCACTTCAAGAGGATCTTGGGA 192148
QY 841 AGTGTTCGAGACCTTTTCTCACTTCACTGATATCAACCCGGAAGACTTCTTGGGACAG 900
DB 192149 AGTGTTCGAGACCTTTTCTCACTTCACTGATATCAACCCGGAAGACTTCTTGGGACAG 192208
QY 901 GTGAAGAAAGATGAGTGTGCTGACCAAGATGCTGTGGAGAACTGCCCCAGGCTGAC 960
DB 192209 GTGAAGAAAGATGAGTGTGCTGACCAAGATGCTGTGGAGAACTGCCCCAGGCTGAC 192268
QY 961 AGGCGAGGCTTACGTAGACAGATGTTATCACTGAGCCCAACTTACTTTAGCAAGGATG 1020

|||||
DB 192269 AGGCGAGGCTTACGTAGACAGATGTTATCACTGAGCCCAACTTACTTTAGCAAGGATG 192328
QY 1021 CTGAACCAAAACCATAGAGTGGACATGCTGATGACAGATGAACACTTCCCATTAAC 1080
DB 192329 CTGAACCAAAACCATAGAGTGGACATGCTGATGACAGATGAACACTTCCCATTAAC 192388
QY 1081 TATTTAGGATGATACCAAGACATCAAGAGAAAGGTGGAGAACTGCTCACTCTAGG 1140
DB 192389 TATTTAGGATGATACCAAGACATCAAGAGAAAGGTGGAGAACTGCTCACTCTAGG 192448
QY 1141 AACTGTAGATGATGAGTGTGAGAGTGTCCAGGCGCTTATGATCATTTTCTCACTGCTG 1200
DB 192449 AACTGTAGATGATGAGTGTGAGAGTGTCCAGGCGCTTATGATCATTTTCTCACTGCTG 192508
QY 1201 GGAACCTCACAAATATCTTCTGCTTCTTGGGGTTCAGCCCAAGCTGTCACAAAATCA 1260
DB 192509 GGAACCTCACAAATATCTTCTGCTTCTTGGGGTTCAGCCCAAGCTGTCACAAAATCA 192568
QY 1261 GATATTTCCCTTATTTCCAGATTTCCAGACATGTCACCCCAATTAATTAACCCCACTT 1320
DB 192569 GATATTTCCCTTATTTCCAGATTTCCAGACATGTCACCCCAATTAATTAACCCCACTT 192628
QY 1321 CAGCCCAATCACGTGGAGAGAGTGAATCTTCCCTTTCTGGAATCTCAAGCATTTACT 1380
DB 192629 CAGCCCAATCACGTGGAGAGAGTGAATCTTCCCTTTCTGGAATCTCAAGCATTTACT 192688
QY 1381 TTCAAGGATCAAAACACGACATTAATGATTAATAACCTTAAAGGCAAAATTTCAAT 1440
DB 192689 TTCAAGGATCAAAACACGACATTAATGATTAATAACCTTAAAGGCAAAATTTCAAT 192747
QY 1441 CTTCCTTCAAGGATGACAGAACTTGGCAAACTCTGAGGCTGTGCAAGAAAGATG 1500
DB 192748 CTTCCTTCAAGGATGACAGAACTTGGCAAACTCTGAGGCTGTGCAAGAAAGATG 192807
QY 1501 TTCAATTTTAAAGATTTGTCTTGGCTGGGTGTGAGGCAATGATCAAGAGATCA 1560
DB 192808 TTCAATTTTAAAGATTTGTCTTGGCTGGGTGTGAGGCAATGATCAAGAGATCA 192867
QY 1561 GGAAGTTGAGACCAACCTGGCCCAATGATGAACCCCAATCTTACCAAAAAAATTCAA 1620
DB 192868 GGAAGTTGAGACCAACCTGGCCCAATGATGAACCCCAATCTTACCAAAAAAATTCAA 192927
QY 1621 ATGAGCGCGCGT 1680
DB 192928 ATGAGCGCGCGT 192987
QY 1681 TTGAACCCAGAGAGT 1740
DB 192988 TTGAACCCAGAGAGT 193047
QY 1741 GGAAGTGAAGCTGTCTCAAAAAAAGATGCTCTCAACTTTGCTTCTCTCACT 1800
DB 193048 GGAAGTGAAGCTGTCTCAAAAAAAGATGCTCTCAACTTTGCTTCTCTCACT 193107
QY 1801 GCAACATTTTGTGATTTGAATGAAGTACCTTCACTTCACTTCACTTCACTTCACTTCACT 1860
DB 193108 GCAACATTTTGTGATTTGAATGAAGTACCTTCACTTCACTTCACTTCACTTCACTTCACT 193167
QY 1861 TCTCACTAGGATGAAGACATTTCTCACTGCTTCTCACTTCACTTCACTTCACTTCACTTCACT 1920
DB 193168 TCTCACTAGGATGAAGACATTTCTCACTGCTTCTCACTTCACTTCACTTCACTTCACTTCACT 193227
QY 1921 GGTCTCACTGT 1980
DB 193228 GGTCTCACTGT 193287
QY 1981 ACACTGAAGAAGTGTCTTCAACACAGCTGAATCAATGTTAATGATTAATAAAT 2040
DB 193288 ACACTGAAGAAGTGTCTTCAACACAGCTGAATCAATGTTAATGATTAATAAAT 193347
QY 2041 CCAGACTACTTCAAGCTTAAATGCTTTTAAATCATTAATAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2100

Db 19348 CCCAGACTACTTACGCTTTATGCTTTATTCATAAAAACTGTAAAGCTAGACTGAA 193407

Qy 2101 CCATTGGAACCTTTAACTCAGACTCTGTGATTCAGAGTCGGGAACCTTAGTTCTATCTG 2160
|||||

Db 193408 CCATTGGAACCTTTAACTCAGACTCTGTGATTCAGAGTCGGGAACCTTAGTTCTATCTG 193467
|||||

Qy 2161 AATCCAGACAGCCACCTTAGTATCTGCCAACTAATGATTAATAAATAC 2216
|||||

Db 193468 AATCCAGACAGCCACCTTAGTATCTGCCAACTAATGATTAATAAATAC 193523
|||||

RESULT 4
AC003025

LOCUS AC003025 112309 bp DNA linear PRI 23-JUL-1998
DEFINITION Human Chromosome 11p12.2 PAC clone PD466a11, complete sequence.
AC003025
VERSION AC003025.1 GI:3337308
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
1 (bases 1 to 112309)
Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,
Buetner,J., Bumeister,R., Card,P., deSallboat,F., Dunn,J.,
English,C., Ehrhridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
HTGS Submission
Unpublished
2 (bases 1 to 112309)
Evans,G.A., Athanasiou,M., Bradbury,P., Brignac,S., Bumeister,R.,
Davis,C., English,C., Franklin,T.L., Garner,H.R., Gordon,M.,
Gotway,G., Grant,O., Hahner,L., Harris,J., Hinson,S.,
Narayanaswamy,U., Newton,J., O'Brien,K., Oliver,T., Patel,P.,
Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R.,
Syed,M., Valenzuela,D., Ward,T. and Wilson,R.
Direct Submission
Submitted (21-OCT-1997) Genome Science and Technology Center,
University of Texas Southwestern Medical Center at Dallas, 5323
Harry Hines Blvd, Dallas, TX 75235-8591, USA
3 (bases 1 to 112309)
Evans,G.A., Athanasiou,M., Aguayo,P., Armstrong,D., Basit,M.,
Buetner,J., Bumeister,R., Card,P., deSallboat,F., Dunn,J.,
English,C., Ehrhridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G.,
Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N.,
Major,T., McFarland,J., Newton,J., Osborne-Lawrence,S.,
Schageman,J., Schultz,R.A., Stimson,S., Syed,M. and Ward,T.
Direct Submission
Submitted (23-JUL-1998) Genome Science and Technology Center,
University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA
On Jul 23, 1998 this sequence version replaced gi:2554967.
IMPORTANT: This submission contains the entire insert of clone
PD466a11. PD466a11 comes from a PAC library constructed at the
Roswell Park Cancer Institute by the Plier de Jong group. This
clone has been finished according to strict quality criteria and
as compressions and repetitive elements. The expected phred/Phrap
calculated errors/10kb is 0.18. In addition, this sequence has
been finished such that 99.9% of consensus base calls consist of
either double-stranded coverage or 2 types of labeling chemistry on
one strand.
Further information regarding the map of this region or
annotation of PD466a11 can be found at
http://gestec.smmcd.edu/chromosome.htm.
CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome
11p12.2 Best's disease region mapped between STS D1S461 and EST
AHNAK. This region spans over 1.5 Mbp.
MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA)
MAILED CLONE OVERLAP: HTGS submitted PAC clones PD051903 and
PD075693.

FEATURES
Source
repeat_region
Location/Qualifiers
1..112309
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
complement(4341..4656)
/rpt_family="Alu"
complement(9880..10120)
/rpt_family="Alu"
10163..10282
/rpt_family="Alu"
complement(11232..12157)
/rpt_family="Alu"
13351..13560
/rpt_family="L1"
complement(15092..15146)
/rpt_family="MIR"
15826..15932
/rpt_family="MIR"
17011..17146
/rpt_family="MIR"
complement(18018..18313)
/rpt_family="Alu"
18493..18741
/rpt_family="Alu"
18820..18954
/rpt_family="Alu"
complement(21374..21661)
/rpt_family="Alu"
22805..23212
/rpt_family="Alu"
23499..23752
/rpt_family="Alu"
24118..24472
/rpt_family="Alu"
25005..25113
/rpt_family="MIR"
31160..31442
/rpt_family="Alu"
complement(32651..32947)
/rpt_family="Alu"
complement(35970..36025)
/rpt_family="MIR"
38958..39145
/rpt_family="Alu"
44878..45480
/rpt_family="Alu"
49769..50071
/rpt_family="Alu"
complement(50172..50310)
/rpt_family="Alu"
50902..50982
/rpt_family="MIR"
51038..51659
/rpt_family="Alu"
54494..54769
/rpt_family="Alu"
55529..55623
/rpt_family="THE1"
complement(55691..55967)
/rpt_family="Alu"
complement(56049..56257)
/rpt_family="Alu"
56721..57100
/rpt_family="THE1"
58402..58680
/rpt_family="Alu"
complement(58748..59345)
/rpt_family="Alu"
59678..59904
/rpt_family="Alu"
61149..61424
/rpt_family="Alu"

repeat_region complement (61930, .62242)
/rpt_family="Alu"
62376, .62710
repeat_region /rpt_family="MLT1"
62915, .63118
repeat_region /rpt_family="Alu"
complement (63349, .63455)
/rpt_family="Alu"
64339, .64614
repeat_region /rpt_family="Alu"
complement (64707, .64984)
/rpt_family="Alu"
repeat_region complement (65596, .65811)
/rpt_family="Alu"
complement (66500, .66753)
/rpt_family="Alu"
repeat_region /rpt_family="Alu"
67311, .67542
repeat_region /rpt_family="Alu"
67835, .68119
repeat_region /rpt_family="Alu"
68860, .69150
repeat_region /rpt_family="Alu"
69164, .69272
repeat_region /rpt_family="Alu"
69354, .70245
repeat_region /rpt_family="Alu"
70655, .70939
repeat_region /rpt_family="Alu"
70933, .71075
repeat_region /rpt_family="L1"
71119, .71395
repeat_region /rpt_family="Alu"
71404, .71530
repeat_region /rpt_family="LTRS"
71536, .71637
repeat_region /rpt_family="LTRS"
72720, .72948
repeat_region /rpt_family="L1"
complement (73488, .73633)
/rpt_family="Alu"
repeat_region complement (73925, .74200)
/rpt_family="Alu"
75721, .75955
repeat_region /rpt_family="Alu"
77185, .77449
repeat_region /rpt_family="Alu"
78093, .78222
repeat_region /rpt_family="L1"
78227, .78294
repeat_region /rpt_family="L1"
complement (80458, .80707)
/rpt_family="Alu"
repeat_region complement (81544, .81974)
/rpt_family="Alu"
82055, .82305
repeat_region /rpt_family="Alu"
82514, .83096
repeat_region /rpt_family="Alu"
complement (83685, .83940)
/rpt_family="Alu"
85336, .85850
repeat_region /rpt_family="Alu"
complement (86429, .86483)
/rpt_family="MIR"
86902, .87183
repeat_region /rpt_family="Alu"
87646, .87792
Query Match 98.7%; Score 2187; DB 9; Length 112309;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2204; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 1 ATCTCGTTTCTTTCAGCGCTGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGG 60

|||||
Db 91454 ATCTCGTTTCTTTCAGCGCTGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGG 91513
QY 61 ACAGAGAGATGCTCAAGCTGGCATCTTGGCCGCTTCTTAGGCTGCAATGCCATGATC 120
Db 91514 ACAGAGAGATGCTCAAGCTGGCATCTTGGCCGCTTCTTAGGCTGCAATGCCATGATC 91573
QY 121 ACCATCCTCCAGAGGCAAACTCAAGGACCAAACTAAGTGGCCCAAGAGGAATCCCTTC 180
Db 91574 ACCATCCTCCAGAGGCAAACTCAAGGACCAAACTAAGTGGCCCAAGAGGAATCCCTTC 91633
QY 181 TCACAGAGGCGCTGCCCAAAAACCAAGGACAGCAAAAGAAAGTTAGGGGCGCAGAGAG 240
Db 91634 TCACAGAGGCGCTGCCCAAAAACCAAGGACAGCAAAAGAAAGTTAGGGGCGCAGAGAG 91693
QY 241 ACAACAAAGGCTGGAAAGCTTAAAGCTGTGAGACGCTTCAAGTCTGCCCATCTATACGA 300
Db 91694 ACAACAAAGGCTGGAAAGCTTAAAGCTGTGAGACGCTTCAAGTCTGCCCATCTATACGA 91753
QY 301 GGCACAGGCTATCAAGTGGCCCAAGAGCGCGCTCAAGCGCCCATCCCATGTTCTGCCCC 360
Db 91754 GGCACAGGCTATCAAGTGGCCCAAGAGCACCCCTCAAGCGCCCATCCCATGTTCTGCCCC 91813
QY 361 TAGAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAAACAAAA 420
Db 91814 TAGAACCATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAAACAAAA 91873
QY 421 GCTTAAAGACTGTGAGTCTGGGCGCAAGAAAGTTGAATGCTCTCAAGAGCATG 480
Db 91874 GCTTAAAGACTGTGAGTCTGGGCGCAAGAAAGTTGAATGCTCTCAAGAGCATG 91933
QY 481 GGGCCTTGAATGAGCAACCAAGATATCTCAAGTGAAGAGAAACGTGAGATTAAAC 540
Db 91934 GGGCCTTGAATGAGCAACCAAGATATCTCAAGTGAAGAGAAACGTGAGATTAAAC 91993
QY 541 TGAAGATATGCGAGAGATCCCGAAATATCAAGTCAAAAGAACTTTGGAAATCAACAA 600
Db 91994 TGAAGATATGCGAGAGATCCCGAAATATCAAGTCAAAAGAACTTTGGAAATCAACAA 92053
QY 601 CCAACATACACTACCTCACTCAAAAGATCAATGATCTTATTTGGGCTTTGGAAAACAGT 660
Db 92054 CCAACATACACTACCTCACTCAAAAGATCAATGATCTTATTTGGGCTTTGGAAAACAGT 92113
QY 661 CTGTCTCCACCGTGAACCAAGGCGACATGTCCTGTCGCGCCACCCAGCTTCCCTTG 720
Db 92114 CTGTCTCCACCGTGAACCAAGGCGACATGTCCTGTCGCGCCACCCAGCTTCCCTTG 92173
QY 721 CTCTAGACCTACCTTCTCTCAAAATTTCTTAGGTTCCATCACTGCCAGAGCACTGG 780
Db 92174 CTCTAGACCTACCTTCTCTCAAAATTTCTTAGGTTCCATCACTGCCAGAGCACTGG 92233
QY 781 ACCTAGCGCCAGCAGCTGGGTTATATCTTGGCCACTTCAAGGATCTTAGGA 840
Db 92234 ACCTAGCGCCAGCAGCTGGGTTATATCTTGGCCACTTCAAGGATCTTAGGA 92293
QY 841 AGTGTTCGGGACCTTTTTCACCTTCACTTCAACCCGGAAGACTTTTGGAGACAG 900
Db 92294 AGTGTTCGGGACCTTTTTCACCTTCACTTCAACCCGGAAGACTTTTGGAGACAG 92353
QY 901 GTGAAGAAAGATGAGTTGTGTCGACCAAGATCTGTGGAGAACTGCCCGGCTGAC 960
Db 92354 GTGAAGAAAGATGAGTTGTGTCGACCAAGATCTGTGGAGAACTGCCCGGCTGAC 92413
QY 961 AGGCGAGGCTTAGCTAGCAGATGTTATCACTGGCCCAACTTACTTTGAGCAAGGTTG 1020
Db 92414 AGGCGAGGCTTAGCTAGCAGATGTTATCACTGGCCCAACTTACTTTGAGCAAGGTTG 92473
QY 1021 CTGAACCAAAACATAGAGTGGAGTCACTGGAATAGATGAACATCTCCCACTAAC 1080
Db 92474 CTGAACCAAAACATAGAGTGGAGTCACTGGAATAGATGAACATCTCCCACTAAC 92533
QY 1081 TATTTAGGTTAGTACCAAGCACTACAGAAAGGTTGGAGAACTGCTCACTCTAGG 1140
|||||

[illegible][illegible]

ORIGIN

Query Match	98.7%	Score 2187	DB 9	Length 163024
Best Local Similarity	99.5%	Pred. No. 0		
Matches 2204; Conservative	0	Matches	11	

QY	1	ATCTCTGTTTCTTTCAGGCTGAACAAGAGGATGGAATTCAGTTCCAGCCCAATCAGAGG	60
Db	124425	ATCTCTGTTTCTTTCAGGCTGAACAAGAGGATGGAATTCAGCCCAATCAGAGG	12448
QY	61	ACGAGAGGATGCTCAGCGTCGATATTGGCGCTTCTTAGGCTGCAAGTCCATATC	120
Db	124485	ACGAGAGGATGCTCAGCGTCGATATTGGCGCTTCTTAGGCTGCAAGTCCATATC	124544
QY	121	ACCATCTCCAGGGCAAACTCAAGGACCAATCTAGTGCCCAAGAGGAATCCCTTC	180
Db	124545	ACCATCTCCAGGGCAAACTCAAGGACCAATCTAGTGCCCAAGAGGAATCCCTTC	124604
QY	181	TCCACGAGGGCTGGCCCAAAAACAAGAGGACCAACAGTAAAGGGCCAGGAAG	240
Db	124605	TCCACGAGGGCTGGCCCAAAAACAAGAGGACCAACAGTAAAGGGCCAGGAAG	124664
QY	241	ACAACAAGGCTGGAAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGA	300
Db	124665	ACAACAAGGCTGGAAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGA	124724
QY	301	GCGCAGGCTACTACAGTGCCTCCACAGAGCGCCCTCAGCCCACTCCCAATGTTCTTCCCC	360
Db	124725	GCGCAGGCTACTACAGTGCCTCCACAGAGCGCCCTCAGCCCACTCCCAATGTTCTTCCCC	124784
QY	361	TAGAACCATCAGCGCGCTCAAAAGTTTCAAGTGTCAAGGATAGACACCAAGACAAA	420
Db	124785	TAGAACCATCAGCGCGCTCAAAAGTTTCAAGTGTCAAGGATAGACACCAAGACAAA	124844
QY	421	GCTTAAAGACTGAGTTCTGGGGCCAAAGAAAGTTTAAATGTTCTTCAAGAGCGATG	480

Db 124845 GCTTAAAGACTGTGAGTCTTGCGGCCAAGAAATTTTGAATTGCTCTCAGAGAGCATG 124904
Qy 481 GGGCTTTGATGAGACACCCGAAAGTATCTCAAGTGAAGAGAAACCTGTGAATTTAAAC 540
Db 124905 GGGCTTTGATGAGACACCCGAAAGTATCTCAAGTGAAGAGAAACCTGTGAATTTAAAC 124964
Qy 541 TGAAGGATATGAGAGATCCCGAAAAATCACCTCAAAAGAACCTTTGGAACAATCAACA 600
Db 124965 TGAAGGATATGAGAGATCCCGAAAAATCACCTCAAAAGAACCTTTGGAACAATCAACA 125024
Qy 601 CCAACAATACACTACACTCACTCAAGATCACTGATGATCTTATTTGGGCTTTGAAAAACAGT 660
Db 125025 CCAACAATACACTACACTCAAGATCACTGATGATCTTATTTGGGCTTTGAAAAACAGT 125084
Qy 661 CTGTCTCTCCACTGGAACAGAGGGGCACTGCAATGCGCTGTGCCCCACCCAGCTTCCCTTG 720
Db 125085 CTGTCTCTCCACTGGAACAGAGGGGCACTGCAATGCGCTGTGCCCCACCCAGCTTCCCTTG 125144
Qy 721 CTCTGAGCCTTACCTTCTCCACAAATTTCTAGGGTTCCATCACTGACAGAGCACTGG 780
Db 125145 CTCTGAGCCTTACCTTCTCCACAAATTTCTAGGGTTCCATCACTGACAGAGCACTGG 125204
Qy 781 ACCTACGCCCAAGCACTGCTGGGGTATATATCTTGGCACTTCAAGGGATCTAGGGA 840
Db 125205 ACCTACGCCCAAGCACTGCTGGGGTATATATCTTGGCACTTCAAGGGATCTAGGGA 125264
Qy 841 AGCTTTCCGGGACCTTTTCTCACTTCACTGCTGTATCACTCCGGAAGACTTTTGGAGCAG 900
Db 125265 AGCTTTCCGGGACCTTTTCTCACTTCACTGCTGTATCACTCCGGAAGACTTTTGGAGCAG 125324
Qy 901 GTGAAGAAAGATAGGTTGTGTGACAGAAATGCTGTGGAGAACTGCCCCAGGGCTGAC 960
Db 125325 GTGAAGAAAGATAGGTTGTGTGACAGAAATGCTGTGGAGAACTGCCCCAGGGCTGAC 125384
Qy 961 AGGCCAGGCTTAGCTGAGCAGATGTTATCACTGGCCCCAATTACTTTAGCAAGGGGTGG 1020
Db 125385 AGGCCAGGCTTAGCTGAGCAGATGTTATCACTGGCCCCAATTACTTTAGCAAGGGGTGG 125444
Qy 1021 CTGACCCCAAAACCATGAGTGGGAGTCACTGATGACAGATGAACACTTCCCCCATTAAC 1080
Db 125445 CTGACCCCAAAACCATGAGTGGGAGTCACTGATGACAGATGAACACTTCCCCCATTAAC 125504
Qy 1081 TATTTAGGGTATGATACCAAGCACTACAGAAAGGGTGGAGAACTGCCCCCATTAAC 1140
Db 125505 TATTTAGGGTATGATACCAAGCACTACAGAAAGGGTGGAGAACTGCCCCCATTAAC 125564
Qy 1141 AACTGTAGATGTGAGTGTGAGGGTGTCCAGGCCCTTAGTCAATTTTCTCACTGCTTG 1200
Db 125565 AACTGTAGATGTGAGTGTGAGGGTGTCCAGGCCCTTAGTCAATTTTCTCACTGCTTG 125624
Qy 1201 GGAACCTCACCAAAATATCTTCTGCTTCTGGGGTCAAGCCCAAGCTGTCAAAAATCA 1260
Db 125625 GGAACCTCACCAAAATATCTTCTGCTTCTGGGGTCAAGCCCAAGCTGTCAAAAATCA 125684
Qy 1261 GATATTTCCCTTATTCAGATTTCTGAGCACTGTCAACCAATTTAAACACCCCACT 1320
Db 125685 GATATTTCCCTTATTCAGATTTCTGAGCACTGTCAACCAATTTAAACACCCCACT 125744
Qy 1321 CAGCCCAATCAAGTGGAGGAAGTGAATCTTCCCTTTCTGAGTCTCAAGCAGTTACT 1380
Db 125745 CAGCCCAATCAAGTGGAGGAAGTGAATCTTCCCTTTCTGAGTCTCAAGCAGTTACT 125804
Qy 1381 TTTCACGGGTGACAAACGCACTTATATGATTGAACCTTTAAAGGCAACAAATTTCAAT 1440
Db 125805 TTTCACGGGTGACAAACGCACTTATATGATTGAACCTTTAAAGGCAACAAATTTCAAT 125863
Qy 1441 CTTCGCTTACGAGTATGACAGAACTTGGCAACATCTGTGGCTGTCAAGCAAGAGATG 1500
Db 125864 CTTCGCTTACGAGTATGACAGAACTTGGCAACATCTGTGGCTGTCAAGCAAGAGATG 125923
Qy 1501 TTTCATATTTAAGAACTTGTCTTGGCTGGGTGTGAGAGCAATCACTACAGAGATCA 1560
Db 125924 TTTCATATTTAAGAACTTGTCTTGGCTGGGTGTGAGAGCAATCACTACAGAGATCA 125983

Qy 1561 GGAGTTTGAACCAACTGGGCCAACAATGATGAACCCCATCTCTACCAAAAAAATACAA 1620
Db 125984 GGAGTTTGAACCAACTGGGCCAACAATGATGAACCCCATCTCTACCAAAAAAATACAA 126043
Qy 1621 ATCAGCTGGCCGTCTGTGTGTGCTGTAGTCCCAAGCAGAGAGTTGAGGGGAAATTCG 1680
Db 126044 ATCAGCTGGCCGTCTGTGTGTGCTGTAGTCCCAAGCAGAGAGTTGAGGGGAAATTCG 126103
Qy 1681 TTGAACCCAGAGAGTGTGTGTGCTGAGTGAATGACAACTGCAATTCAGCTGGGGCAG 1740
Db 126104 TTGAACCCAGAGAGTGTGTGTGCTGAGTGAATGACAACTGCAATTCAGCTGGGGCAG 126163
Qy 1741 GGAGTGAACCTGCTCAAAAAAATGATGCTCAACCTTTGCCCCCTACT 1800
Db 126164 GGAGTGAACCTGCTCAAAAAAATGATGCTCAACCTTTGCCCCCTACT 126223
Qy 1801 GCACAATTTTGGTATTTGAATGAAGTACCTTCCATCTTATGCTTTATCTTTCAAT 1860
Db 126224 GCACAATTTTGGTATTTGAATGAAGTACCTTCCATCTTATGCTTTATCTTTCAAT 126283
Qy 1861 TCTCACTAGGAGTGAAGCACTTCTTACCTGCTTCTTAATGGGATGCTTGGCCAGCA 1920
Db 126284 TCTCACTAGGAGTGAAGCACTTCTTACCTGCTTCTTAATGGGATGCTTGGCCAGCA 126343
Qy 1921 GGTCTCACCTGTGTGTACCAAGCAGAGCACTGATTCAGTCAAGCCATACAGCTGTCC 1980
Db 126344 GGTCTCACCTGTGTGTACCAAGCAGAGCACTGATTCAGTCAAGCCATACAGCTGTCC 126403
Qy 1981 ACACGTAAAGACGTGTCTTCAACAACAGCTGAATCAATAGTTAGTTAATAGATAAAT 2040
Db 126404 ACACGTAAAGACGTGTCTTCAACAACAGCTGAATCAATAGTTAATAGATAAAT 126463
Qy 2041 CCCAGACTACTTCAAGCTTTAATGCTTTTATTCATPAAAACTGTGAAGCTNAGCTGAA 2100
Db 126464 CCCAGACTACTTCAAGCTTTAATGCTTTTATTCATPAAAACTGTGAAGCTNAGCTGAA 126523
Qy 2101 CCATTGGAAACATTTAATCTCAGCTGTGATTCAGAGTGGGGAAACCTTAGTCTATCTG 2160
Db 126524 CCATTGGAAACATTTAATCTCAGCTGTGATTCAGAGTGGGGAAACCTTAGTCTATCTG 126583
Qy 2161 AATCCAAAGCAGCAGCACTTAGTATAGTCCCAACTAATAGTTAATTAATAC 2216
Db 126584 AATCCAAAGCAGCAGCACTTAGTATAGTCCCAACTAATAGTTAATTAATAC 126639

RESULT 6
AC087451/c 163915 bp DNA 1linear HTG 04-JULY-2001
LOCUS AC087451/1
DEFINITION Homo sapiens chromosome 11 clone RP11-810P12 map 11, WORKING DRAFT
SEQUENCE, 7 unordered pieces.
ACCESSION AC087451
VERSION AC087451.2 GI:14595836
HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 163915)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Baer, V., Boguski, L., Bouckge, B., Brown, A.,
Cambridge, J., Campione, A., Chao, L., Collins, S., Collins, S.,
Collins, A., Cooke, P., DeBartolo, K., Dewar, K., Diaz, J.,
Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardner, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heath, A., Horton, L., Hume, W., Iliev, I., Johnson, R.,
Jones, C., Karas, A., LaRocque, K., Lamaz, R., Landers, T.,
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Margie, N., Mathew, C., McCarthy, M., McEwan, P., McKernan, K.,

Dp	150666	ACCTACGCCAGAC	CTGCGCTTGGGGATAT	TACTTGGCCAC	CTTCACAGGAT	CTCTAGGGA	150667
Qy	841	AGTGTTCGGGAC	CTTTTCTCACTTCA	CCCTGGTATCA	CCCGGAAG	CTTCTTGGGAC	900
Dp	150606	AGTGTTCGGGAC	CTTTTCTCACTTCA	CCCTGGTATCA	CCCGGAAG	CTTCTTGGGAC	900
Qy	901	GTGAAGGAAG	TAGAGTGTGGCTG	ACAGAAATGCTG	CTGGAGAA	CTAGCCCGAGGCTG	960
Dp	150546	GTGAAGGAAG	TAGAGTGTGGCTG	ACAGAAATGCTG	CTGGAGAA	CTAGCCCGAGGCTG	960
Qy	961	AGGCCAGGCTT	AGCTGAGCAGAT	GTATTCATCTG	GGCCCCCA	CTTACTTTAGCAAGGGTGG	1020
Dp	150486	AGGCCAGGCTT	AGCTGAGCAGAT	GTATTCATCTG	GGCCCCCA	CTTACTTTAGCAAGGGTGG	1020
Qy	1021	CTGACCCAAA	CCATGATGAGTGG	CGACGTCA	CGCTGGATG	ACAGTACACATTTCCCCCA	1080
Dp	150426	CTGACCCAAA	CCATGATGAGTGG	CGACGTCA	CGCTGGATG	ACAGTACACATTTCCCCCA	1080
Qy	1081	TATTTAAGGT	TACTACCCAGC	CTACAGAAAAGGGTGG	AGAAATGCTG	CTCACTCTTAG	1140
Dp	150366	TATTTAAGGT	TACTACCCAGC	CTACAGAAAAGGGTGG	AGAAATGCTG	CTCACTCTTAG	1140
Qy	1141	AACCTGGTAG	TAGTGTGAGTGTG	AGGGGTGC	AGCGCCCTTAG	GCATTTTCTCACTGCTG	1200
Dp	150306	AACCTGGTAG	TAGTGTGAGTGTG	AGGGGTGC	AGCGCCCTTAG	GCATTTTCTCACTGCTG	1200
Qy	1201	GGAACTCTCA	CAAAATACTCTT	CTTGCTTCTT	GGGGGTCA	GGCCCAAGCTGTCA	1260
Dp	150246	GGAACTCTCA	CAAAATACTCTT	CTTGCTTCTT	GGGGGTCA	GGCCCAAGCTGTCA	1260
Qy	1261	GATATTTCCCT	TATTCAGATTTT	CCTGGACACTG	CACCCAA	TTATTAACCCCACTT	1320
Dp	150186	GATATTTCCCT	TATTCAGATTTT	CCTGGACACTG	CACCCAA	TTATTAACCCCACTT	1320
Qy	1321	CAGCCCCAA	TCACTGTGGAGAG	AGGTAACTTCC	CTTTCTGGAT	CTCAGAGCTTACT	1380
Dp	150126	CAGCCCCAA	TCACTGTGGAGAG	AGGTAACTTCC	CTTTCTGGAT	CTCAGAGCTTACT	1380
Qy	1381	TTCAAGGGT	CAGAACACG	CAGCTATTTAT	GTAAACCTT	TAAGGGCAACATTTCA	1440
Dp	150066	TTCAAGGGT	CAGAACACG	CAGCTATTTAT	GTAAACCTT	TAAGGGCAACATTTCA	1440
Qy	1441	CTTGCTTCT	AGCTTAAGACAG	GAATCTTG	CAAAATCTTG	GGCTGTTCAGCAAAAGATG	1500
Dp	150007	CTTGCTTCT	AGCTTAAGACAG	GAATCTTG	CAAAATCTTG	GGCTGTTCAGCAAAAGATG	1500
Qy	1501	TTCAATTTAA	AAATCTGTCTT	GGGCGTGGG	GTGGAGGCA	AGTGAATCA	1560
Dp	149947	TTCAATTTAA	AAATCTGTCTT	GGGCGTGGG	GTGGAGGCA	AGTGAATCA	1560
Qy	1561	GGAGTTTGA	CAACCTTG	GGCAACATG	ATGTAACCCCA	CTCTCA	1620
Dp	149887	GGAGTTTGA	CAACCTTG	GGCAACATG	ATGTAACCCCA	CTCTCA	1620
Qy	1621	ATCAGCTGG	CGTGTGTGTG	CTGTAGTCC	CAACGCA	AGGTTAGGGGAGAA	1680
Dp	149827	ATCAGCTGG	CGTGTGTGTG	CTGTAGTCC	CAACGCA	AGGTTAGGGGAGAA	1680
Qy	1681	TTGAACCC	AGAGGTGTGTG	CTGAGTGA	TTAGCA	CTGCAATCC	1740
Dp	149767	TTGAACCC	AGAGGTGTGTG	CTGAGTGA	TTAGCA	CTGCAATCC	1740
Qy	1741	GGAGTGA	CACTGTCTCA	AAAAAAGAA	ATGCTCTCA	ACCTTGGACCTCTACT	1800
Dp	149707	GGAGTGA	CACTGTCTCTCA	AAAAAAGAA	ATGCTCTCA	ACCTTGGACCTCTACT	1800
Qy	1801	GCACATTTG	GTATTTGA	ATGAAGTA	CTTCACTTA	ATGCTGTTAATACTT	1860
Dp	149647	GCACATTTG	GTATTTGA	ATGAAGTA	CTTCACTTA	ATGCTGTTAATACTT	1860
Qy	1861	TCTCACT	GTGAAGTGA	GCATTTCT	CAATCTG	CTTAATGGGAGTCTTGG	1920
Dp	149587	TCTCACT	GTGAAGTGA	GCATTTCTCA	ATCTGCTTCA	TATGGGAGTGTGCTTGG	1920

Qy	1921	GATCTCCACCTGTGTGTACACAGAGGACACTGATCCAGTACAGCCCTTAACGCTGTCC	1980
Db	149527	GGTCTCCACCTGTGTGTACACAGAGGACACTGATCCAGTACAGCCCTTAACGCTGTCC	1494688
Qy	1981	ACACTGGAAGACGTGTCTTACAAACGCGTGAATCAAATGTTAGCTTAATAGATAAAAT	2040
Db	149467	ACACTGGAAGACATGTCTTACAAACGCGTGAATCAAATGTTAGCTTAATAGATAAAAT	1494080
Qy	2041	CCCAAGACTACTTAAGCCTTTAATGCTTTTATTCATTAATAAACTGTGAAAGCTTGAATCGAA	2100
Db	149407	CCCAAGACTACTTAAGCCTTTAATGCTTTTATTCATTAATAAACTGTGAAAGCTTGAATCGAA	1493488
Qy	2101	CCATTGGAAACATTTAATCTCAGACTCTTGATTCAGAGTCGGAAACCTTAGTTCTATCTG	2160
Db	149347	CCATTGGAAACATTTAATCTCAGACTCTTGATTCAGAGTCGGAAACCTTAGTTCTATCTG	1492888
Qy	2161	AATCCAAAGACGCCACACCTTAGTATAGTACGCCAAACCTAATGAGTTAATAAATC	2216
Db	149287	AATCCAAAGACGCCACACCTTAGTATAGTACGCCAAACCTAATGAGTTAATAAATC	149232

RESULT 7

Locus AP003733 166867 bp DNA linear PRI 17-JUN-2003

DEFINITION Homo sapiens genomic DNA, chromosome 11 clone:RP11-810P12, complete genome.

ACCESSION AP003733

VERSION AP003733.5 GI:31790740

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
1 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens genomic DNA
Published Only in Database (2001)
2 (bases 1 to 166867)
Hattori M., Ishii K., Toyoda A., Taylor T.D., Hong-Seog P.,
Fujiyama A., Yada T., Totoki Y., Watanabe H. and Sakaki Y.
Direct Submission
Submitted (11-JUN-2001) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Saitiro-chou,Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:hattorice@cc.riken.go.jp, URL:http://hngp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
On Jun 16, 2003 this sequence version replaced gi:17426128.

```
source
1.166867
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q"
/clone="RP11-810P12"
```

ORIGIN

Query Match	98.7%	Score 2187	DB 9	Length 166867
Best Local Similarity	99.5%	Pred. No. 0		
Matches 2204	Conservative	0	Mismatches 11	Indels 1
			Gaps	1
QY	1	ATCTCTGTTTTCTTTTCAGCCTGGAACAAAGAGAGATGGAGTTCCAGCCCAATCAGAGAG	60	
Db	12470	ATCTCTGTTTTCTTTTCAGCCTGGAACAAAGAGAGATGGAGTTCCAGCCCAATCAGAGAG	12529	
QY	61	ACGAGAGAGATGCTACGCTGGGATCATTTGGCGGCTTCTTAGGCTGCGAGTGCATGATC	120	
Db	12530	ACGAGAGAGATGCTACGCTGGGATCATTTGGCGGCTTCTTAGGCTGCGAGTGCATGATC	125689	
QY	121	ACCATCTTCCACAGGGAACCTCAGAGCCAAACTACTGTGGCCCAAGAGAGATCCCTTC	180	

Dh 12590 ACCATCCGCCAGGCGAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTC 12649
Qy 181 TCCACGAGGGCTGCGCCAAAACCAAGGACAGCAAGAAACGTTGGGGCCGAGAG 240
Db 12550 TCCACGAGGGCTGCGCCAAAACCAAGGACAGCAAGAAACGTTGGGGCCGAGAG 12709
Qy 241 ACAACAGGCGCTGAAGCTTAAAGCTGTGAGAGCCCTTCAAGCTGCGCCACTGTATCAGA 300
Db 12710 ACAACAGGCGCTGAAGCTTAAAGCTGTGAGAGCCCTTCAAGCTGCGCCACTGTATCAGA 12769
Qy 301 GGCAGGCTTACTACAGTGGCCCAAGAGCCCGCTTCAAGCTGCGCCACTGTATCAGA 12769
Db 12770 GGCAGGCTTACTACAGTGGCCCAAGAGCCCGCTTCAAGCTGCGCCACTGTATCAGA 12829
Qy 361 TGAACCATCAGCGCGCTCAAGCTTCAAGTGCACAGGATGACACCAAGACAAA 420
Db 12830 TGAACCATCAGCGCGCTCAAGCTTCAAGTGCACAGGATGACACCAAGACAAA 12889
Qy 421 GCTTAAAGCTGTGAGTCTGGGGCAAGAAAGTTTGAATTGCTCTCAGAGAGGATG 480
Db 12890 GCTTAAAGCTGTGAGTCTGGGGCAAGAAAGTTTGAATTGCTCTCAGAGAGGATG 12949
Qy 481 GGGCCTTATGAGACCCAGAAATCTCAAGTGGAGAGAGAAACTGTGAGTTTAAAC 540
Db 12950 GGGCCTTATGAGACCCAGAAATCTCAAGTGGAGAGAGAAACTGTGAGTTTAAAC 13009
Qy 541 TGACGGATATGCCAGAGATCCCGAAATCACTCAAGAAAGCTTGGAAACATGACCA 600
Db 13010 TGACGGATATGCCAGAGATCCCGAAATCACTCAAGAAAGCTTGGAAACATGACCA 13069
Qy 601 CCAACATACACACTCACTCAAGATCACTGATGCTTATGGGCTTGGAAACAGGT 660
Db 13070 CCAACATACACACTCACTCAAGATCACTGATGCTTATGGGCTTGGAAACAGGT 13129
Qy 661 CTGTCCCTCACTCAAGAGGCGGCACTGATGCTTATGGGCTTGGAAACAGGT 720
Db 13130 CTGTCCCTCACTCAAGAGGCGGCACTGATGCTTATGGGCTTGGAAACAGGT 13189
Qy 721 CTGTGAGCTTACCTTCTCACAATTTCTTGAAGGTTCCATCACTGCGACAGCACTGG 780
Db 13190 CTGTGAGCTTACCTTCTCACAATTTCTTGAAGGTTCCATCACTGCGACAGCACTGG 13249
Qy 781 ACCTAGCGCCAGACCTGGGTTGGGATATATCTTGGCCACCTTCAAGAGGATCTTGAAG 840
Db 13250 ACCTAGCGCCAGACCTGGGTTGGGATATATCTTGGCCACCTTCAAGAGGATCTTGAAG 13309
Qy 841 AGTGTTCGGGACCTTTTCTCACTTCACTGCTGTATCACTGGGAAAGCTTGGGACAG 900
Db 13310 AGTGTTCGGGACCTTTTCTCACTTCACTGCTGTATCACTGGGAAAGCTTGGGACAG 13369
Qy 901 GTGAAGGAAGATGAGTGTGCTGACCAAGATGCTGCTGAGAACTGCCCCAGGCTGAC 960
Db 13370 GTGAAGGAAGATGAGTGTGCTGACCAAGATGCTGCTGAGAACTGCCCCAGGCTGAC 13429
Qy 961 AGGCGAGGCTTAACTGAGAGATGTTATCACTGGCCCAACTTACTTGAAGAAAGGTTG 1020
Db 13430 AGGCGAGGCTTAACTGAGAGATGTTATCACTGGCCCAACTTACTTGAAGAAAGGTTG 13489
Qy 1021 CTGACCCAAAACCATGAGTGGGAGTCACTGTGATGACAGATGAACACTTCCCCCAATAC 1080
Db 13490 CTGACCCAAAACCATGAGTGGGAGTCACTGTGATGACAGATGAACACTTCCCCCAATAC 13549
Qy 1081 TATTTAGGGTATGATCCCAAGCACTACAGAAAGGGTGGAGAACTGCTCACTCTTAG 1140
Db 13550 TATTTAGGGTATGATCCCAAGCACTACAGAAAGGGTGGAGAACTGCTCACTCTTAG 13609
Qy 1141 AACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db 13610 AACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 13669
Qy 1201 GGAACCTCACCAAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1260
Db 13670 GGAACCTCACCAAAATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 13729

Qy 1261 GATATTTCCCTTATTTCCAGATTTTCTGAGACCTGTGACCCCAATTAACAACCCCACTT 1320
Db 13730 GATATTTCCCTTATTTCCAGATTTTCTGAGACCTGTGACCCCAATTAACAACCCCACTT 13789
Qy 1321 CAGCCCAATACCTGGAGAGAGTGTATCTTCCCTTCTTCTGATCTTCAAGCATTTACT 1380
Db 13790 CAGCCCAATACCTGGAGAGAGTGTATCTTCCCTTCTTCTGATCTTCAAGCATTTACT 13849
Qy 1381 TTCAAGGCTCAGAACAGGCTATTAATGTTGAAACTTTAAAGGCAACATTTCAAT 1440
Db 13850 TTCAAGGCTCAGAACAGGCTATTAATGTTGAAACTTTAAAGGCAACATTTCAAT 13908
Qy 1441 CTTCCTTCTGAGCTTAAGACAGAACTTGGCAAAACATCTGTGGCTGTTCAAGAAAGATG 1500
Db 13909 CTTCCTTCTGAGCTTAAGACAGAACTTGGCAAAACATCTGTGGCTGTTCAAGAAAGATG 13968
Qy 1501 TTCAATTTTAAAGATCTTGTCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1560
Db 13969 TTCAATTTTAAAGATCTTGTCTTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 14028
Qy 1561 GAGGTTTGAACCAACCTGGCCCAATGATGAAACCCCATCTTCAACCAAAATTAACA 1620
Db 14029 GAGGTTTGAACCAACCTGGCCCAATGATGAAACCCCATCTTCAACCAAAATTAACA 14088
Qy 1621 ATCAGCTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1680
Db 14089 ATCAGCTGGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 14148
Qy 1681 TTGAACCCAGAGAGT 1740
Db 14149 TTGAACCCAGAGAGT 14208
Qy 1741 GAGGTGAGACGTCTCAAAAAAAGAGATGCTCTCAACTTTGGCTCTACT 1800
Db 14209 GAGGTGAGACGTCTCAAAAAAAGAGATGCTCTCAACTTTGGCTCTACT 14268
Qy 1801 GCAACATTTTGTGTTTGAATGAAGTACTTCCATCTTATGCTGTTAATTAATTTCAAT 1860
Db 14269 GCAACATTTTGTGTTTGAATGAAGTACTTCCATCTTATGCTGTTAATTAATTTCAAT 14328
Qy 1861 TCTCATTAGGAGTGAAGACATTTCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
Db 14329 TCTCATTAGGAGTGAAGACATTTCCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 14388
Qy 1921 GGTCTCACCCTGT 1980
Db 14389 GGTCTCACCCTGT 14448
Qy 1981 ACACTGAAGAGTGTCTCTCAAGAGCTGAACTGAATGTTAGCTTAATGAATTAAT 2040
Db 14449 ACACTGAAGAGTGTCTCTCAAGAGCTGAACTGAATGTTAGCTTAATGAATTAAT 2040
Qy 2041 CCGAGCTACTTCAAGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 14508
Db 14509 CCGAGCTACTTCAAGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 14568
Qy 2101 CCATTGAAACATTTAACTGAGCTGATTAAGTGGGAACTTAAATGCTTAAATGCT 2160
Db 14569 CCATTGAAACATTTAACTGAGCTGATTAAGTGGGAACTTAAATGCTTAAATGCT 14628
Qy 2161 AATCCAGAGAGCCACCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 2216
Db 14629 AATCCAGAGAGCCACCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 14684

RESULT 8
AC084857
LOCUS
DEFINITION Homo sapiens chromosome 11 clone CTD-3231N5 map 11, WORKING DRAFT
ACCESSION AC084857
VERSION AC084857.1 GI:11276215

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS 1 (bases 1 to 133683)

TITLE Birren,B., Linton,L., Nuebaum,C., and Lander,E.

JOURNAL Homo sapiens chromosome 11, clone CTD-3231N5

REFERENCE Unpublished

AUTHORS 2 (bases 1 to 133683)

Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N., Anderson,S., Barna,N., Bastien,V., Bede,F., Bogunlavkiy,L., Boukhalter,B., Brown,A., Buckett,G., Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArliano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Headford,A., Horton,L., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Larocque,K., Lamares,R., Landers,T., Lehocsky,J., Levine,R., Lieu,C., Liu,G., Macdonald,P., Margis,N., McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Rieback,M., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Sougnez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strause,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Titrill,A., Travers,M., Trigglio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zaimoun,J., Zimmer,A., and Zody,M.

Direct Submission

Submitted (32-NOV-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smith,A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L11038

Center clone name: 3231.N.5

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 126761 bases at least Q40

Consensus quality: 129952 bases at least Q30

Consensus quality: 131351 bases at least Q20

Insert size: 128000; agarose-fp

Insert size: 132483; sum-of-coverage

Quality coverage: 6.1 in Q20 bases; sum-of-coverage

Quality coverage: 5.9 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently

* consists of 13 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 1459: gap of 1459 bp in length

* 1460 1559: gap of 100 bp

* 1560 2520: contig of 961 bp in length

* 2521 2620: gap of 100 bp

* 2621 6334: contig of 3714 bp in length

* 6335 6434: gap of 100 bp

* 6435 11913: contig of 5479 bp in length

* 11914 12013: gap of 100 bp

* 12014 20756: contig of 8743 bp in length

20757 20856: gap of 100 bp

* 20857 29735: contig of 8879 bp in length

* 29736 29835: gap of 100 bp

* 29836 38645: contig of 8810 bp in length

* 38646 38745: gap of 100 bp

* 38746 52193: contig of 13448 bp in length

* 52194 52293: gap of 100 bp

* 52294 65209: contig of 12916 bp in length

* 65210 65310: gap of 100 bp

* 65310 79796: contig of 14467 bp in length

* 79797 79896: gap of 100 bp

* 79897 95897: contig of 16000 bp in length

* 95897 95996: gap of 100 bp

* 95997 113419: contig of 17423 bp in length

* 113420 113519: gap of 100 bp

* 113520 113683: contig of 20164 bp in length.

Location/Qualifiers

1. 133683

source

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="11"

/map="11"

/clone="CTD-3231N5"

/clone_11b="CITD Human BAC"

1. 1459

misc_feature

/note="assembly_fragment"

1560.2520

misc_feature

/note="assembly_fragment"

2621.6334

misc_feature

/note="assembly_fragment"

6435.11913

misc_feature

/note="assembly_fragment"

12014.20756

misc_feature

/note="assembly_fragment"

20857.29735

misc_feature

/note="assembly_fragment"

29836.38645

misc_feature

/note="assembly_fragment"

38746.52193

misc_feature

/note="assembly_fragment"

52294.65209

misc_feature

/note="assembly_fragment"

65310.79796

misc_feature

/note="assembly_fragment"

79897.95896

misc_feature

/note="assembly_fragment"

95997.113419

misc_feature

/note="assembly_fragment"

113520.113683

misc_feature

/note="assembly_fragment"

ORIGIN

Query Match 98.6%; Score 2185.4; DB 2; Length 133683;

Best Local Similarity 99.4%; Pred. No. 0;

Matches 2203; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1 ATCTCTGTTTCTTTCACCTGTAACAAGAGAGATGAGTTCAGCCCAATCAGAGG 60

Db 92296 ATCTCTGTTTCTTTCACCTGTAACAAGAGAGATGAGTTCAGCCCAATCAGAGG 92355

QY 61 ACAGAGAGATGTCACGCTGGATCATTTGCCCTTCTAGGCCGACAGTCCCATGATC 120

Db 92356 ACAGAGAGATGTCACGCTGGATCATTTGCCCTTCTAGGCCGACAGTCCCATGATC 92415

QY 121 ACCATCTCCAGGAGCAACTCAGAGCAAACTACTGTGSCCAAGAGGAATCCCTTC 180

Db 92416 ACCATCTCCAGGAGCAAACTCAGAGCAAACTACTGTGSCCAAGAGGAATCCCTTC 92475

QY 181 TCCAGAGGAGGCTGCCCAAAAACACAGAGGACCAAGAAAGTTAGGGGCGCAGAGG 240

Db 92476 TCCAGAGGAGGCTGCCCAAAAACACAGAGGACCAAGAAAGTTAGGGGCGCAGAGG 92535

Qy	241	ACAAACAAGCCTGGAAGCTTAAAGCTGTGAAGCCCTCAAGTCTGCCCACTGTATAGA	300
Db	92536	ACAAACAAGCCTGGAAGCTTAAAGCTGTGAAGCCCTCAAGTCTGCCCACTGTATAGA	92596
Qy	301	GGCCAGGCTACTACAGTGGCCCAAGAGGCCCTTCAGGCCCACTCCATTTCTTCCCC	360
Db	92596	GGCCAGGCTACTACAGTGGCCCAAGAGGCCCTTCAGGCCCACTCCATTTCTTCCCC	92655
Qy	361	TGAAACCATCAGGCCCTTCAAGCTTCAAGTGCACAGGAGATMGACCAAGACAAA	420
Db	92656	TGAAACCATCAGGCCCTTCAAGCTTCAAGTGCACAGGAGATMGACCAAGACAAA	92715
Qy	421	GCTTAAAGCTGTGAGTTCGTGGGCCAAAGAAAGTTTGAATTGCTCAGAGAGGATG	480
Db	92716	GCTTAAAGCTGTGAGTTCGTGGGCCAAAGAAAGTTTGAATTGCTCAGAGAGGATG	92775
Qy	481	GGGCGCTTGATGAGACCCCAAGAGTATCTCAAGTGAAGAGAAACCTGTGGAGTTAAC	540
Db	92776	GGGCGCTTGATGAGACCCCAAGAGTATCTCAAGTGAAGAGAAACCTGTGGAGTTAAC	92835
Qy	541	TGACGGATATGCGAGAGATCCCGGAAATCACTCAAGAACCTTTGGACAACTCA	600
Db	92836	TGACGGATATGCGAGAGATCCCGGAAATCACTCAAGAACCTTTGGACAACTCA	92895
Qy	601	CCAAACATACACTTAATCTAAAGATCAGATGATCTTAAATGGGCGTTGGAAAAAGGT	660
Db	92896	CCAAACATACACTTAATCTAAAGATCAGATGATCTTAAATGGGCGTTGGAAAAAGGT	92955
Qy	661	CTGTCTCCACCTGAACCAAGGGGCACTGATTTGCCCTGTGCCCAACCCAGTCTCC	720
Db	92956	CTGTCTCCACCTGAACCAAGGGGCACTGATTTGCCCTGTGCCCAACCCAGTCTCC	93015
Qy	721	CTCTAGGCTTACCTCTCTCTCAAAATTTCTTGAAGTTTCATCTCTGCAAGAGCACTGG	780
Db	93016	CTCTAGGCTTACCTCTCTCTCAAAATTTCTTGAAGTTTCATCTCTGCAAGAGCACTGG	93075
Qy	781	ACCTACGCCGCCAGCACTGGCTTGGGGTATATCTTGGCACCTTCACAGAGATCTTAA	840
Db	93076	ACCTACGCCGCCAGCACTGGCTTGGGGTATATCTTGGCACCTTCACAGAGATCTTAA	93135
Qy	841	AGTGTGGGACCTTTTCTCACTTCAACCTGTGATCAACCCGMAAGCTTCTTGGGACAG	900
Db	93136	AGTGTGGGACCTTTTCTCACTTCAACCTGTGATCAACCCGMAAGCTTCTTGGGACAG	93195
Qy	901	GTGAAGAGATGAGTGTGTCTGACCAAAATCTGTCTGAGAACTGCCCAAGGCTGAC	960
Db	93196	GTGAAGAGATGAGTGTGTCTGACCAAAATCTGTCTGAGAACTGCCCAAGGCTGAC	93255
Qy	961	AGGCCAGGCTTACTGACGAGTGTATATCACTGAGGCCCAACTTATCTTGAACAAAGG	1020
Db	93256	AGGCCAGGCTTACTGACGAGTGTATATCACTGAGGCCCAACTTATCTTGAACAAAGG	93315
Qy	1021	CTGACCCAAACCATGAGGTGGCAGTCACTGAGTGAATGAACACTTCCCCATAAC	1080
Db	93316	CTGACCCAAACCATGAGGTGGCAGTCACTGAGTGAATGAACACTTCCCCATAAC	93375
Qy	1081	TATTTAAGGCTACTACCCAAAGCACTACAGAAAGGGTGGAGAACTGSCCTCATCTAG	1140
Db	93376	TATTTAAGGCTACTACCCAAAGCACTACAGAAAGGGTGGAGAACTGSCCTCATCTAG	93435
Qy	1141	AACTGTAGATGTGAGGTGTGAGGTTGCCAGGCGCTTGAAGTCAATTTTCTCACTGCC	1200
Db	93436	AACTGTAGATGTGAGGTGTGAGGTTGCCAGGCGCTTGAAGTCAATTTTCTCACTGCC	93495
Qy	1201	GGAACTTACAAATATCTTCTGTCTCTTGTGGGGTCAAGCCCAAGGCTGTCAAAATCA	1260
Db	93496	GGAACTTACAAATATCTTCTGTCTCTTGTGGGGTCAAGCCCAAGGCTGTCAAAATCA	93555
Qy	1261	GATATTTCCCTTATTCAGATTTCCCTGACACTGTGCACTCCCAATATTAACACCCCACTT	1320
Db	93556	GATATTTCCCTTATTCAGATTTCCCTGACACTGTGCACTCCCAATATTAACACCCCACTT	93615
Qy	1321	CAGCCCAATCAGGTGGAGAAATGTAACTTCCCTTTCTGAGATTCACAGCACTTACT	1380

[illegible]

AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 160169)
AUTHORS Waterston,R.H.
JOURNAL Direct Submission
Submitted (15-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Apr 23, 2000 this sequence version replaced gi:5754984.

COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0810P12
----- Summary Statistics -----
Sequencing vector: M13, 100%
Sequencing vector: plasmid: 0%
Chemistry: Dye-primer ET, 100% of reads
Chemistry: Dye-terminator Big Dye, 0% of reads
Assembly program: Phrap; Version 0.990319
Consensus quality: 147316 bases at least Q40
Consensus quality: 151042 bases at least Q30
Consensus quality: 152875 bases at least Q20
Insert size: 165000; agarose-fp
Insert size: 157869; sum-of-contigs
Quality coverage: 3.91 in Q20 bases; sum-of-contigs
Quality coverage: 4.17 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1926: contig of 1926 bp in length
* 1927 2026: gap of unknown length
* 2027 4202: contig of 2176 bp in length
* 4203 4302: gap of unknown length
* 4303 7163: contig of 2861 bp in length
* 7164 7263: gap of unknown length
* 7264 9961: contig of 2698 bp in length
* 9962 10061: gap of unknown length
* 10062 12575: contig of 2514 bp in length
* 12576 12675: gap of unknown length
* 12676 15545: contig of 2870 bp in length
* 15546 15645: gap of unknown length
* 15646 20941: contig of 5296 bp in length
* 20942 21041: gap of unknown length
* 21042 23966: contig of 2925 bp in length
* 23967 24066: gap of unknown length
* 24067 27231: contig of 3165 bp in length
* 27232 27331: gap of unknown length
* 27332 30550: contig of 3219 bp in length
* 30551 30650: gap of unknown length
* 30651 35202: contig of 4552 bp in length
* 35203 35302: gap of unknown length
* 35303 40201: contig of 4899 bp in length
* 40202 40301: gap of unknown length
* 40302 45758: contig of 5457 bp in length
* 45759 45858: gap of unknown length
* 45859 51249: contig of 5391 bp in length
* 51250 51349: gap of unknown length
* 51350 58267: contig of 6918 bp in length
* 58268 58367: gap of unknown length
* 58368 66916: contig of 8549 bp in length
* 66917 77145: contig of 10129 bp in length
* 77146 77245: gap of unknown length

* 77246 86493: contig of 9248 bp in length
* 86494 86593: gap of unknown length
* 86594 96688: contig of 10095 bp in length
* 96689 96788: gap of unknown length
* 96789 107169: contig of 10381 bp in length
* 107170 107269: gap of unknown length
* 107270 117047: contig of 9778 bp in length
* 117048 117147: gap of unknown length
* 117148 127120: contig of 9973 bp in length
* 127121 127220: gap of unknown length
* 127221 142029: contig of 14809 bp in length
* 142030 142129: gap of unknown length
* 142130 160169: contig of 18040 bp in length.
* 160169 160169: contig of 18040 bp in length.
----- Location/Qualifiers -----
1. 160169
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/clone="RP11-810P12"
1. 1926
/note="assembly_name:Contig18"
2027. 4202
/note="assembly_name:Contig19"
4303. 7163
/note="assembly_name:Contig20
clone_end:"7"
vector_side:"left"
7264. 9961
/note="assembly_name:Contig21"
10062. 12575
/note="assembly_name:Contig22"
12676. 15545
/note="assembly_name:Contig23"
15646. 20941
/note="assembly_name:Contig24"
21042. 23966
/note="assembly_name:Contig25"
24067. 27231
/note="assembly_name:Contig26"
27332. 30550
/note="assembly_name:Contig27"
30651. 35202
/note="assembly_name:Contig28"
35303. 40201
/note="assembly_name:Contig29"
40302. 45758
/note="assembly_name:Contig30"
45859. 51249
/note="assembly_name:Contig31"
51350. 58267
/note="assembly_name:Contig32"
58368. 66916
/note="assembly_name:Contig33"
67017. 77145
/note="assembly_name:Contig34"
77246. 86493
/note="assembly_name:Contig35"
86594. 96688
/note="assembly_name:Contig36"
96789. 107169
/note="assembly_name:Contig37"
107270. 117047
/note="assembly_name:Contig38"
117148. 127120
/note="assembly_name:Contig39
clone_end:SP6
vector_side:"right"
127221. 142029
/note="assembly_name:Contig40"
142130. 160169
/note="assembly_name:Contig41"

ORIGIN

Query Match 72.6%; Score 1608.4; DB 2; Length 160169;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1626; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 577 AAGAACCTTGGAAACATCATCCAAACATACATCACTCACTCAAGATCATATGATC 636
 DB 51139 AAGACNTTGGAAACATCATCCAAACATACATCACTCAAGATCATATGATC 51080
 QY 637 CTATTTGGGCGCTTGGAAAAAGAGTCTGTCTCTCACTCAAGATCATATGATC 636
 DB 51079 CTATTTGGGCGCTTGGAAAAAGAGTCTGTCTCTCACTCAAGATCATATGATC 51020
 QY 697 TGTGCCCCACCCAGCTTCCCTTGTCTGAGCTTACCTTCTCAATTTCTTAAGGT 756
 DB 51019 TGTGCCCCACCCAGCTTCCCTTGTCTGAGCTTACCTTCTCAATTTCTTAAGGT 50960
 QY 757 TCCATCATGCGCAGAGCACTGGAACCTACGCCAGCATGAGCTTGGGGATATATCTGG 816
 DB 50959 TCCATCATGCGCAGAGCACTGGAACCTACGCCAGCATGAGCTTGGGGATATATCTGG 50900
 QY 817 CCACCTTCAAGAGGATCTTAGAGAAAGTTCGGGACCTTTCTCACTTCACTCGATATC 876
 DB 50899 CCACCTTCAAGAGGATCTTAGAGAAAGTTCGGGACCTTTCTCACTTCACTCGATATC 50840
 QY 877 ACCCGGAAGACTTCTTGGGACACAGTGAAGAAAGATGAGTGTGTGACAGAAATGCTG 936
 DB 50839 ACCCGGAAGACTTCTTGGGACACAGTGAAGAAAGATGAGTGTGTGACAGAAATGCTG 50780
 QY 937 CTGAGAACTGCCCCAGGGGTGACAGAGGCTTAGCTGAGCAATATATCTAGCTGCC 996
 DB 50779 CTGAGAACTGCCCCAGGGGTGACAGAGGCTTAGCTGAGCAATATATCTAGCTGCC 50720
 QY 997 CCAACTTACTTTAGAGAAAGGTGCTGACCCAAACCATGAGTGGCAGTGAAGTGTG 1056
 DB 50719 CCAACTTACTTTAGAGAAAGGTGCTGACCCAAACCATGAGTGGCAGTGAAGTGTG 50660
 QY 1057 ACGATGAACACTTCCCATTAATATTTAGGGTGTACCCCAACCTACAGGAAAGGT 1116
 DB 50659 ACGATGAACACTTCCCATTAATATTTAGGGTGTACCCCAACCTACAGGAAAGGT 50600
 QY 1117 GGCAGGAAGCTGCTCACTCTAGAAACCTGTAAGTGTGAGTGTGAGGCTGACAGGCC 1176
 DB 50599 GGCAGGAAGCTGCTCACTCTAGAAACCTGTAAGTGTGAGTGTGAGGCTGACAGGCC 50540
 QY 1177 CTTAGGTCAATTTCTCACTGCTGAGAACTCACCAAAATATCTTCTTGGGGT 1236
 DB 50539 CTTAGGTCAATTTCTCACTGCTGAGAACTCACCAAAATATCTTCTTGGGGT 50480
 QY 1237 CAGCCAAAGCTGTCACAAATGAGATTTCCCTTATTCAGATTTCTGACACTGT 1296
 DB 50479 CAGCCAAAGCTGTCACAAATGAGATTTCCCTTATTCAGATTTCTGACACTGT 50420
 QY 1297 CACCAATTTAAACACCCCACTTACGCCCAATCACTGAGGAGGATGTAATCTTCCCT 1356
 DB 50419 CACCAATTTAAACACCCCACTTACGCCCAATCACTGAGGAGGATGTAATCTTCCCT 50360
 QY 1357 TTTCTGATCTCAAGAGTACTTTCACGGGTGAGAACCGCACTTATGATTGAA 1416
 DB 50359 TTTCTGATCTCAAGAGTACTTTCACGGGTGAGAACCGCACTTATGATTGAA 50301
 QY 1417 CCTTAAAGGGCAAAATTCATCTTGTCTTGAAGTGAAGCAAGGAACTTGGCAAT 1476
 DB 50300 ACTTAAAGGGCAAAATTCATCTTGTCTTGAAGTGAAGCAAGGAACTTGGCAAT 50241
 QY 1477 CTGTGCGCTGTCAAGAAAGATGTCATATTTAAGATCTTGTGTGGCTGGGTGTGG 1536
 DB 50240 CTGTGCGCTGTCAAGAAAGATGTCATATTTAAGATCTTGTGTGGCTGGGTGTGG 50181
 QY 1537 AGGCAAGTGAATCAGAGAGGTGAGAGTGTGAGAACCACTGGCCAACTATGAAAC 1596
 DB 50180 AGGCAAGTGAATCAGAGAGGTGAGAGTGTGAGAACCACTGGCCAACTATGAAAC 50121

QY 1597 CCATCTTACCAAAAAAATATCAATATAGCTGGCCGTGTGTGTCTGTAGTCCCAAC 1656
 DB 50120 CCATCTTACCAAAAAAATATCAATATAGCTGGCCGTGTGTGTCTGTAGTCCCAAC 50061
 QY 1657 GCAGAGGTTGAGGGAGAAATGCTTGAACCCAGAGAGTGTGTGTGAGATGAG 1716
 DB 50060 GCAGAGGTTGAGGGAGAAATGCTTGAACCCAGAGAGTGTGTGTGAGATGAG 50001
 QY 1717 CAATGCAATCCAGCCCTGGGCGACGAGTGAAGTGTCTCAAAAAAAGAT 1776
 DB 50000 CAATGCAATCCAGCCCTGGGCGACGAGTGAAGTGTCTCAAAAAAAGAT 49941
 QY 1777 GGTCTCAACTTTTCCCTCTCACTGCAATTTTGTGATTTGAATGAAGTACCTTCA 1836
 DB 49940 GGTCTCAACTTTTCCCTCTCACTGCAATTTTGTGATTTGAATGAAGTACCTTCA 49881
 QY 1837 TACTATGCTGTAAATATCTTCAATTTCTCACTTGGAGTGAAGCAATCTTACCTGCTTC 1896
 DB 49880 TACTATGCTGTAAATATCTTCAATTTCTCACTTGGAGTGAAGCAATCTTACCTGCTTC 49821
 QY 1897 CTATTTGGGATCTTTCGCCAGCAGAGTCTCACTGTGTGACACGAGCACTGAT 1956
 DB 49820 CTATTTGGGATCTTTCGCCAGCAGAGTCTCACTGTGTGACACGAGCACTGAT 49761
 QY 1957 CCAGTCAAGCATACAGCTGTCTCCACACTGAAGACGTGTCTTCAACACGCTGAATCA 2016
 DB 49760 CCAGTCAAGCATACAGCTGTCTCCACACTGAAGACGTGTCTTCAACACGCTGAATCA 49701
 QY 2017 ATGTTTACTTAAATATGATTAATAATCCAGACTTCACTTCACTTATGCTTATATCAT 2076
 DB 49700 ATGTTTACTTAAATATGATTAATAATCCAGACTTCACTTCACTTATGCTTATATCAT 49641
 QY 2077 AAAAATCTGTAAGACTGACGTGAACCACTTGAACCTTTAATCACTGATCTGATTCAGA 2136
 DB 49640 AAAAATCTGTAAGACTGACGTGAACCACTTGAACCTTTAATCACTGATCTGATTCAGA 49581
 QY 2137 GTGGGAAACCTTATGTTTATCTGAATTCAGACAGCCACACTTATGATTCAGCCAA 2196
 DB 49580 GTGGGAAACCTTATGTTTATCTGAATTCAGACAGCCACACTTATGATTCAGCCAA 49521
 QY 2197 CTATGAGTTTATTAATATC 2216
 DB 49520 CTATGAGTTTATTAATATC 49501

RESULT 10
 AC004588/c
 LOCUS AC004588
 DEFINITION Homo sapiens, *** SEQUENCING IN PROGRESS ***, 19 unordered pieces.
 AC004588.2
 VERSION AC004588.2 GI:21406530
 KEYWORDS HTG: HTGS PHASE1
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 106648)
 Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basil, M.,
 Beutner, J., Bumeister, R., Card, P., desailloat, C., Dunn, J.,
 English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G.,
 Grant, O., Hahner, L., Harris, U., Lewis, E., Loo, H., Loo, K.N.,
 McFarland, J., Miller, R., Newton, J., Osborne-Lawrence, S.,
 Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.
 HTGS Submission
 Unpublished
 2 (bases 1 to 106648)
 Evans, G.A., Athanasiou, M., Aguayo, P., Armstrong, D., Basil, M.,
 Beutner, J., Bumeister, R., Card, P., desailloat, C., Dunn, J.,
 English, C., Ethridge, S., Garner, H.R., Gee, V., Gordon, M., Gotway, G.,
 Grant, O., Hahner, L., Harris, U., Lewis, E., Loo, H., Loo, K.N.,
 McFarland, J., Miller, R., Newton, J., Osborne-Lawrence, S.,
 Schageman, J., Schultz, R.A., Stimson, S., Syed, M. and Ward, T.
 Direct Submission

TITLE

JOURNAL

Submitted (18-APR-1998) Genome Science & Technology Center,
University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA

COMMENT

On Jun 13, 2002 this sequence version replaced gi:3063499.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
* 1
* 2483: contig of 2483 bp in length
* 2484 2583: gap of 100 bp
* 2584 4622: contig of 2039 bp in length
* 4623 4722: gap of 100 bp
* 4723 6725: contig of 2003 bp in length
* 6726 6825: gap of 100 bp
* 6826 8448: contig of 1623 bp in length
* 8449 8548: gap of 100 bp
* 8549 11050: contig of 2502 bp in length
* 11051 11150: gap of 100 bp
* 11151 13554: contig of 2404 bp in length
* 13555 13654: gap of 100 bp
* 13655 17064: contig of 3410 bp in length
* 17065 17164: gap of 100 bp
* 17165 20800: contig of 3636 bp in length
* 20801 20900: gap of 100 bp
* 20901 23967: contig of 3067 bp in length
* 23968 24067: gap of 100 bp
* 24068 26216: contig of 2149 bp in length
* 26217 26316: gap of 100 bp
* 26317 31202: contig of 4886 bp in length
* 31203 31302: gap of 100 bp
* 31303 35714: contig of 4412 bp in length
* 35715 35814: gap of 100 bp
* 35815 39320: contig of 3506 bp in length
* 39321 39420: gap of 100 bp
* 39421 47809: contig of 8389 bp in length
* 47810 47909: gap of 100 bp
* 47910 55372: contig of 7463 bp in length
* 55373 55472: gap of 100 bp
* 55473 56654: contig of 1182 bp in length
* 56655 56755: gap of 100 bp
* 56755 65820: contig of 9066 bp in length
* 65821 65920: gap of 100 bp
* 65921 87612: contig of 21692 bp in length
* 87613 87712: gap of 100 bp
* 87713 106648: contig of 18936 bp in length.
```

FEATURES

source

```
1..106648
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1..104848
/notes="CHROMOSOMAL LOCUS: This PAC clone comes from the  
Chromosome -  
11p12.2 Best's disease region mapped between STS D11S461  
and EST AHNAK. This region spans over 1.5 Mbp.  
MARKER CONFIRMATION: EST; FTH (ferritin heavy chain mRNA)
```

misc_feature

```
repeat_region
473..745
/rpt_family="Alu"
complement(1443..1649)
repeat_region
/rpt_family="MER21"
complement(2690..2959)
repeat_region
/rpt_family="Alu"
3427..3725
/rpt_family="Alu"
complement(5231..5605)
repeat_region
/rpt_family="LTR3"
complement(6706..6996)
```

```
/rpt_family="Alu"
7018..7459
/rpt_family="MER21"
7067..7420
/rpt_family="MER39"
complement(7570..7856)
/rpt_family="Alu"
11234..11801
/rpt_family="Alu"
complement(12283..12373)
/rpt_family="MIR"
12666..13031
/rpt_family="Alu"
13546..14276
/rpt_family="Alu"
complement(14646..14917)
/rpt_family="Alu"
complement(15040..15200)
/rpt_family="Alu"
15289..15571
/rpt_family="Alu"
16315..16437
/rpt_family="Alu"
complement(16471..16746)
/rpt_family="Alu"
complement(16780..16914)
/rpt_family="MER21"
16905..17072
/product="Human ferritin heavy chain mRNA"
17331..17503
/rpt_family="Alu"
complement(17959..18124)
/rpt_family="MER21"
18889..19131
/rpt_family="Alu"
23180..23501
/rpt_family="Alu"
complement(23733..24024)
/rpt_family="Alu"
complement(24646..24717)
/rpt_family="MIR"
24775..25060
/rpt_family="Alu"
26119..26366
/rpt_family="Alu"
complement(26555..26616)
/rpt_family="Alu"
26767..27818
/rpt_family="Alu"
complement(28535..28821)
/rpt_family="Alu"
28895..29109
/rpt_family="Alu"
29148..29439
/rpt_family="Alu"
31140..31278
/rpt_family="MER4"
31282..31675
/rpt_family="LTR7"
31793..32411
/rpt_family="Alu"
complement(32903..33202)
/rpt_family="Alu"
complement(33214..33708)
/rpt_family="LTR8"
complement(33709..33804)
/rpt_family="LTR9"
33826..34102
/rpt_family="Alu"
complement(34112..34172)
/rpt_family="LTR9"
34470..34614
/rpt_family="Alu"
```

repeat_region 34913. .38086
/rpt_family="L1"
repeat_region 34994. .35249
/rpt_family="MER25"
repeat_region 36889. .37180
/rpt_family="Alu"
repeat_region 38416. .38688
/rpt_family="Alu"
repeat_region 38864. .39464
/rpt_family="Alu"
repeat_region 39952
complement(39826.
/rpt_family="Alu"
repeat_region 40060. .40131
complement(40264.
/rpt_family="Alu"
repeat_region 40575. .40614
/rpt_family="Mariner1"
repeat_region 41866. .42139
/rpt_family="Alu"
repeat_region 42633. .42885
/rpt_family="Alu"
repeat_region 42948. .43175
/rpt_family="Alu"
repeat_region 43360. .43497
/rpt_family="Alu"
repeat_region 43744. .43924
/rpt_family="MLT1"

Query Match 70.5%; Score 1561.6; DB 2; Length 106648;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1578; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 628 ACATGATCCTTATGAGGCTTGGAAACAAGTCTGTCTCCACTGAACCAAGGGGCACT 687
Db 106648 ACATGATCCTTATGAGGCTTGGAAACAAGTCTGTCTCCACTGAACCAAGGGGCACT 106589
QY 688 GCATTGCGCTGTGCCCCACCCAGCTTCCCTGTCTGAGGCTTACCTCTCCCAATT 747
Db 106588 GCATTGCGCTGTGCCCCACCCAGCTTCCCTGTCTGAGGCTTACCTCTCCCAATT 106529
QY 748 TCTAGAGGTTCCATCATGCGCAAGACCAACATGACCTTACGCGCAAGCACTGCTTGGGTA 807
Db 106528 TCTAGAGGTTCCATCATGCGCAAGACCAACATGACCTTACGCGCAAGCACTGCTTGGGTA 106469
QY 808 TATACCTGGCCACTTCAACAGGATCTTAGGGAAGTGTGGGACCTTTCTCACTTAC 867
Db 106468 TATACCTGGCCACTTCAACAGGATCTTAGGGAAGTGTGGGACCTTTCTCACTTAC 106409
QY 868 CTTGATACCCCGGAACCTTTTGGGACCAAGTGAAGAGATGAGTTGTGTGACC 927
Db 106408 CTTGATACCCCGGAACCTTTTGGGACCAAGTGAAGAGATGAGTTGTGTGACC 106349
QY 928 AGAATGCTGTGAGAACTGCCCAAGGCTGACAGGCGAGGCTTACTGACGATGTTA 987
Db 106348 AGAATGCTGTGAGAACTGCCCAAGGCTGACAGGCGAGGCTTACTGACGATGTTA 106289
QY 988 TCACTGGCCCAACTTACTTTAGCAAGAGGTGTGACCAAAACCATGAGTGTGCACTC 1047
Db 106288 TCACTGGCCCAACTTACTTTAGCAAGAGGTGTGACCAAAACCATGAGTGTGCACTC 106229
QY 1048 AGCTGATGACAGATGAACTTCCCAATTAATTTAGGGTATGACCAAGCACTTACA 1107
Db 106228 AGCTGATGACAGATGAACTTCCCAATTAATTTAGGGTATGACCAAGCACTTACA 106169
QY 1108 GGAAGAGGTGGAGAACTGCTCACTCTAGAACTGTGATAGTGTGAGGTTG 1167
Db 106168 GGAAGAGGTGGAGAACTGCTCACTCTAGAACTGTGATAGTGTGAGGTTG 106109
QY 1168 TCCAGGCGCTTATGATCTTTCTCACTGCTGGGAACTTCACTTCTTCTT 1227
Db 106108 TCCAGGCGCTTATGATCTTTCTCACTGCTGGGAACTTCACTTCTTCTT 106049

QY 1228 CCTGGGGGTCAGCCCAAGCTGTCACAAAATCAGATATTTCCCTTATTCAGATTTCT 1287
Db 106048 CCTGGGGTCAGCCCAAGCTGTCACAAAATCAGATATTTCCCTTATTCAGATTTCT 105989
QY 1288 GACACTGTCAACCAATTTAAACACCCCACTTACGCCCCAATCAGTGTGAGGAAATGT 1347
Db 105988 GACACTGTCAACCAATTTAAACACCCCACTTACGCCCCAATCAGTGTGAGGAAATGT 105929
QY 1348 AACTTCCCTTTCTGATTTCTCAAGCATTTCTTACGCGGTCAAGACGACTATTA 1407
Db 105928 AACTTCCCTTTCTGATTTCTCAAGCATTTCTTACGCGGTCAAGACGACTATTA 105870
QY 1408 TGATGAAACCTTAAAGGCAACAATTCATCTGCTTCAAGCTAAGACAGACTT 1467
Db 105869 TGATGAAACCTTAAAGGCAACAATTCATCTGCTTCAAGCTAAGACAGACTT 105810
QY 1468 GGCAAAATCTGTGGCTCTTTCAGCAAGAGATGTCAATTTAAGATCTTGTGGGC 1527
Db 105809 GGCAAAATCTGTGGCTCTTTCAGCAAGAGATGTCAATTTAAGATCTTGTGGGC 105750
QY 1528 TGGGTGTGAGGCAAGTGAATCAAGAGGTGAGAGTTTGAACCACTGGCCAACT 1587
Db 105749 TGGGTGTGAGGCAAGTGAATCAAGAGGTGAGAGTTTGAACCACTGGCCAACT 105690
QY 1588 GATGAAACCCCATCTCTACCAAAAAATACAAATCAGCTGGCGTGTGGCTGT 1647
Db 105689 GATGAAACCCCATCTCTACCAAAAAATACAAATCAGCTGGCGTGTGGCTGT 105630
QY 1648 AGTCCCAACGCGAGAGGTGAGGGGAGATTTCTTGAACCAAGAGTGTGGTGCAGT 1707
Db 105629 AGTCCCAACGCGAGAGGTGAGGGGAGATTTCTTGAACCAAGAGTGTGGTGCAGT 105570
QY 1708 GAGATTGACCACTGCAATCCAGCTGGCGACGAGTGTGAGCTGTCAAAAAAATA 1767
Db 105569 GAGATTGACCACTGCAATCCAGCTGGCGACGAGTGTGAGCTGTCAAAAAAATA 105510
QY 1768 AAAAAGATCGTCTCAACCTTTGGCCCTCACTGTGCAACATTTGTGTAATGAAG 1827
Db 105509 AAAAAGATCGTCTCAACCTTTGGCCCTCACTGTGCAACATTTGTGTAATGAAG 105450
QY 1828 TACCTTCATACCTTATCTGTATATCTTCACTTCACTGAGGATGAACATTTCTTA 1887
Db 105449 TACCTTCATACCTTATCTGTATATCTTCACTTCACTGAGGATGAACATTTCTTA 105390
QY 1888 ACCGTCTTCCATTAAGGGGAGTGTGCGCAGCCAGGCTCTCAGCTGTGTGACCAAGAG 1947
Db 105389 ACCGTCTTCCATTAAGGGGAGTGTGCGCAGCCAGGCTCTCAGCTGTGTGACCAAGAG 105330
QY 1948 GACACTGATCAGTCAAGCAATACAGCTGTCCACCTGAGAGAGAGTGTCTTCAACAGC 2007
Db 105329 GACACTGATCAGTCAAGCAATACAGCTGTCCACCTGAGAGAGAGTGTCTTCAACAGC 105270
QY 2008 CTGAATCAATAGTGTGCTTAATATATTAATAAATCCAGACTACTTACGCTTATGCT 2067
Db 105269 CTGAATCAATAGTGTGCTTAATATATTAATAAATCCAGACTACTTACGCTTATGCT 105210
QY 2068 TTTATTCATTAATACTGTGAAGAGTGAAGTGAACATTTGAATTAATCTAGACTT 2127
Db 105209 TTTATTCATTAATACTGTGAAGAGTGAAGTGAACATTTGAATTAATCTAGACTT 105150
QY 2128 GATTCAGAGTGGGAACTTATGTTATCTATGTAATCAAGACGACCACTTATGATA 2187
Db 105149 GATTCAGAGTGGGAACTTATGTTATCTATGTAATCAAGACGACCACTTATGATA 105090
QY 2188 CTGCCAAACTAATGAGTTTAATAATAC 2216
Db 105089 CTGCCAAACTAATGAGTTTAATAATAC 105061

RESULT 11
BC041664 2435 bp mRNA linear PRI 07-OCT-2003
LOCUS BC041664
DEFINITION Homo sapiens vitelliform macular dystrophy (Best disease,

QY 601 CCAACATACACTACACTCAAGATGATGATCTTATTTGGGCTTGAAAAACAGT 660
|
|
|
DB 617 CCAACATACACTACACTCAAGATGATGATCTTATTTGGGCTTGAAAAACAGT 676
|
|
|
QY 661 CTGTCTCTCACTGACCGAGGAGGAGCTGCA 690
|
|
|
DB 677 CTGTCTCTCACTGACCGAGGAGGAGCTGCA 706
|
|
|
RESULT 13
AY357926S10 726 bp DNA linear PRI 26-AUG-2003
LOCUS AY357926S10 726 bp DNA linear PRI 26-AUG-2003
DEFINITION Macaca fascicularis beestrophin (VMD2) gene, exon 10.
ACCESSION AY357935
VERSION AY357935.1 GI:34013779
KEYWORDS
SEGMENT
SOURCE
ORGANISM
Macaca fascicularis (crab-eating macaque)
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopitheciinae; Macaca.
REFERENCE
AUTHORS
Okamoto, H., Umeda, S., Suzuki, M.T., Yoshikawa, Y., Tanaka, Y. and
Iwata, T.
TITLE
Molecular Cloning of VMD2 Gene from Cynomolgus Monkey (Macaca
fascicularis)
JOURNAL
Unpublished
REFERENCE
AUTHORS
Okamoto, H., Umeda, S., Suzuki, M.T., Yoshikawa, Y., Tanaka, Y. and
Iwata, T.
TITLE
Direct Submission
JOURNAL
Submitted (01-AUG-2003) National Institute of Sensory Organs,
National Tokyo Medical Center, 2-5-1, Higashi-gaoka, Meguro-Ku,
Tokyo 152-0021, Japan
FEATURES
source
1..726
/organism="Macaca fascicularis"
/mol_type="genomic DNA"
/db_xref="taxon:9541"
36..674
/gene="VMD2"
/number=10
exon
ORIGIN
Query Match 29.9%; Score 662; DB 9; Length 726;
Best Local Similarity 95.8%; Pred. No. 2.1e-193;
Matches 680; Conservative 0; Mismatches 30; Indels 0; Gaps 0;
QY 1 ATCTCCGTTCTTTCCAGGCTGAACAAGAGGATGAGTCCAGCCCAATCAGAGG 60
|
|
|
DB 17 ATCTCCGTTCTTTCCAGGCTGAACAAGAGGATGAGTCCAGCCCAATCAGAGG 76
|
|
|
QY 61 ACAGAGAGATGCTCAAGCTGATCATTTGGCGCTTCTAGGCGCTGCACTCCATGATC 120
|
|
|
DB 77 ACAGAGAGAGCTCAAGCTGATCATTTGGCGCTTCTAGGCGCTGCACTCCATGATC 136
|
|
|
QY 121 ACCATCTCCAGGAGCAATCAAGACCAACTACTGTGGCCCAAGAGGAATCCCTTC 180
|
|
|
DB 137 ACCATCTCCAGGAGCAATCAAGACCAACTACTGTGGCCCAAGAGGAATCCCTTC 196
|
|
|
QY 181 TCCAGAGGGGCTGCCAANAACCAAGGAGGAGCAAGAAAGTTAGGGGCGCAGGAAG 240
|
|
|
DB 197 TCCAGAGGGGCTGCCAANAACCAAGGAGGAGCAAGAAAGTTAGGGGCGCAGGAAG 256
|
|
|
QY 241 ACAACAAGGCTGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGA 300
|
|
|
DB 257 ACAACAAGGCTGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGA 316
|
|
|
QY 301 GGCAGAGCTTACTAGTGGCCCAAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCC 360
|
|
|
DB 317 GGCAGAGCTTACTAGTGGCCCAAGAGCGCCCTCAGGCCCACTCCCATGTTCTTCCCC 376
|
|
|

QY 361 TAGAACATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGACAAA 420
|
|
|
DB 377 CAGAACATCAGCGCGCTCAAAAGCTTCAAGTGTCAAGGATAGACACCAAGACAAA 436
|
|
|
QY 421 GCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGCGATG 480
|
|
|
DB 437 GCTTAAAGACTGTGAGTTCTGGGGCCAAAGAAAGTTTGAATTGCTCTCAGAGCGATG 496
|
|
|
QY 481 GGGCTTGAAGGAGCAACCAAGATCTCAAGTGAAGGAGAAACCTGTGAGTTTAAAC 540
|
|
|
DB 497 GGGCTTGAAGGAGCAACCAAGATCTCAAGTGAAGGAGAAACCTGTGAGTTTAAAC 556
|
|
|
QY 541 TGAAGATATGCAAGATCCCGAAATACCTCAAGAACTTTGGAACAAACACCA 600
|
|
|
DB 557 TGAAGATATGCAAGATCCCGAAATACCTCAAGAACTTTGGAACAAACACCA 616
|
|
|
QY 601 CCAACATACACTACACTCAAGATGATGATCTTATTTGGGCTTGAAAAACAGT 660
|
|
|
DB 617 CCAACATACACTACACTCAAGATGATGATCTTATTTGGGCTTGAAAAACAGT 676
|
|
|
QY 661 CTGTCTCTCACTGACCGAGGAGGAGGAGCTGCACTTGGGCTTGAAAAACAGT 710
|
|
|
DB 677 CTGTCTCTCACTGACCGAGGAGGAGGAGCTGCACTTGGGCTTGAAAAACAGT 726
|
|
|
RESULT 14
BD1326S9 1263 bp DNA linear PAT 18-SEP-2002
LOCUS BD1326S9 1263 bp DNA linear PAT 18-SEP-2002
DEFINITION Secreted proteins and polynucleotides encoding them.
ACCESSION BD1326S9
VERSION BD1326S9.1 GI:23227604
KEYWORDS
JP 2002504822-A/2
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
Jacobs, K., McCoy, J.M., Lavallie, E.R., Racie, L.A., Treac, M.,
Spaulding, V., Agostino, M.J., Howes, S.H. and Fechtel, K.
TITLE
Secreted proteins and polynucleotides encoding them
JOURNAL
Patent: JP 2002504822-A 2 12-FEB-2002;
GENETICS INSTITUTE INC
PM JP 2002504822-A/2
COMMENT
PD 12-FEB-2002
PR 08-JUN-1998 JP 199503038
PR 11-JUN-1997 US 08/873218, 05-JUN-1998 US 09/092722 PI
KENNETH JACOB, JOHN M MCCOY, EDWARD R LAVALLIE, LISA A RACIE, PI
MAURICE TREACY,
PI VIKKI SPAULDING, MICHAEL J AGOSTINO, STEVEN H HOWES, KIM FECHTEL,
PC C12N15/12, C07K14/47, A61K38/17
CC Strandedness: Double;
CC Topology: Linear;
FH Key
FEATURES
source
1..1263
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 29.1%; Score 643.8; DB 6; Length 1263;
Best Local Similarity 99.7%; Pred. No. 1e-187;
Matches 645; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 17 CAGCTGAACCAAGAGAGATGAGTTCAGGCCCAATCAGAGAGCAGAGAGATGCTCA 76
|
|
|
DB 225 CAGCTGAACCAAGAGAGATGAGTTCAGGCCCAATCAGAGAGCAGAGAGATGCTCA 284
|
|
|
QY 77 CGTGGCATCATTTGGCGCTTCTTCTAGGCTGTGAGTCCCATGATCAGATCTTCCAGGGC 136
|
|
|
DB 285 CGTGGCATCATTTGGCGCTTCTTCTAGGCTGTGAGTCCCATGATCAGATCTTCCAGGGC 344
|
|
|
QY 137 AAACCTCAAGAGCAACCAACTACTGTGCGCCAAAGAGGAATCCCTTCTCAGAGAGGCGCTGCC 196
|
|
|

Db 345 AAACCTCAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGGCCCTGCC 404
QY 197 CAAAACCAACAGGACAGCCAAACAGAACTTAAAGGAGCAGAGAAACAAGGCTTGA 256
Db 405 CAAAACCAACAGGACAGCCAAACAGAACTTAAAGGAGCAGAGAAACAAGGCTTGA 464
QY 257 GCTTAAAGCTGTGAGGAGCTTCAAGTGTGCCCCCACTGTATCAAGGCAAGCTTACAG 316
Db 465 GCTTAAAGCTGTGAGGAGCTTCAAGTGTGCCCCCACTGTATCAAGGCAAGCTTACAG 524
QY 317 TGCCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTGAACCAATGAGGCG 376
Db 525 TGCCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTGAACCAATGAGGCG 584
QY 377 GTCAAGGCTTCAACAGTGTCAAGGCAATGACCAACCAAGAAAGCTTAAAGACTGTGAG 436
Db 585 GTCAAGGCTTCAACAGTGTCAAGGCAATGACCAACCAAGAAAGCTTAAAGACTGTGAG 644
QY 437 TTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGCGAATGGGGCTTGAAGAGA 496
Db 645 TTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGCGAATGGGGCTTGAAGAGA 704
QY 497 CCCAAGAGTATCTCAAGTGTGAGAGAAACCTGTGAGTTAACTGACGATATGCCAGA 556
Db 705 CCCAAGAGTATCTCAAGTGTGAGAGAAACCTGTGAGTTAACTGACGATATGCCAGA 764
QY 557 GATCCCCGAAATTCACCTCAAGAAACCTTTGGAACAATCACCAACCAATTCACATCAC 616
Db 765 GATCCCCGAAATTCACCTCAAGAAACCTTTGGAACAATCACCAACCAATTCACATCAC 824
QY 617 ACTCAAGATCAATGATCTTATTGGGCTTGGAAAAAGTCTG 663
Db 825 ACTCAAGATCAATGATCTTATTGGGCTTGGAAAAAGGATG 871

RESULT 15
AF057169 2210 bp mRNA linear PRI 17-Oct-1998
LOCUS Homo sapiens beetrophin (VMD2) mRNA, alternatively spliced product,
DEFINITION complete cds.
ACCESSION AF057169
VERSION AF057169.1 GI:3335158
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2210)
Petrushkin, K., Koistinen, M.J., Bakall, B., Li, W., Xie, G., Marknell, T.,
Sandgren, O., Forsman, K., Holmgren, G., Andreasson, S., Vujic, M.,
Metzger, M.L., Caskey, C.T. and Madelin, C.
Identification of the gene responsible for Best macular dystrophy
Nat. Genet. 19 (3), 241-247 (1998)
TITLE
JOURNAL MEDLINE
PUBMED 96324772
REFERENCE 2 (bases 1 to 2210)
Petrushkin, K.
AUTHORS Direct Submission
JOURNAL Submitted (03-APR-1998) Human Genetics, Merck Research
Laboratories, West Point, PA 19486, USA
FEATURES
Source
1. .2210
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="11"
1. .2210
/gene="VMD2"
105. .1862
/gene="VMD2"

ORIGIN

Query Match 29.1%; Score 643.8; DB 9; Length 2210;
Best Local Similarity 99.7%; Pred. No. 1.1e-187;
Matches 645; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

/note="Best's macular dystrophy gene; BMD, alternatively
splice isoform"
/codon_start=1
/product="beetrophin"
/protein_id="AAC64343.1"
/db_xref="GI:3335159"
/translation="MTITVTSOVANRILGFSRLLICWBGSIYKLYGFLPLFLCTY
IIRPIRLATTEOOLMPEKTLTYCDSYIQLPISFVLPFTVLTVMWQYENLPW
PDRILMSVGFVSGKDEOGRILRLIRLANIGNVILSVSTAVYKRPSPAOHLVOA
GMPTRPBNKOLEKLSLPHNMFVPMVPMFANLNMKMLGRIIDPILLOSLNEMTLR
TQGGHLYADWISIPVYQVYVAVSFLLICVGRIPKADVPGEHLDVVPV
TLPQFFVYGMLEKVAEQLINPGEEDDDDEFTWMIYDRNLQVSLAVDEHODLPMPER
DMYMNKDEPQPYTAASQFRASFMWGSFTNLSIKKEWEPQPNODEDEDAAGITGR
FLGLSDHHPRPANRRTLLMPKPSLHBEGLPKNHKAAQNVGQEDNKKLAV
DAKPSALYORBGVYSAPOTPLSPFTWFFPLSPASVLSHVSYGLDIDKSLKTVSSG
AKSPFLBSDSGALMHPVSVQVRKTFVFNLTMPETIPENHMLEPLEOSPTNIHTT
LKHMDPVMALENRDEAHS"

QY 17 CAGCTGGAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 76
Db 1202 CAGCTGGAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 1261
QY 77 CGCTGCATCATTTGGCGCGCTTCTAGGCGCTGAGCCCATATCACTCTCCAGAGGC 136
Db 1262 CGCTGCATCATTTGGCGCGCTTCTAGGCGCTGAGCCCATATCACTCTCCAGAGGC 1321
QY 137 AAACCTCAAGACCAAACTGCTGTGCGCCCAAGAGGAATCCCTTCCAGAGGCGCTGCC 196
Db 1322 AAACCTCAAGACCAAACTGCTGTGCGCCCAAGAGGAATCCCTTCCAGAGGCGCTGCC 1381
QY 197 CAAAACCAACAGGACAGCCAAACAGAACTTAAAGGAGCCAGAGAAACAACAAGGCTTGA 256
Db 1382 CAAAACCAACAGGACAGCCAAACAGAACTTAAAGGAGCCAGAGAAACAACAAGGCTTGA 1441
QY 257 GCTTAAAGCTGTGAGAGCGCTTCAAGTGTGCGCCCATGTATCAGAGGCAAGCTTACAG 316
Db 1442 GCTTAAAGCTGTGAGAGCGCTTCAAGTGTGCGCCCATGTATCAGAGGCAAGCTTACAG 1501
QY 317 TGCCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTGAACCAATCAGGCGC 376
Db 1502 TGCCCCCAGAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCTTGAACCAATCAGGCGC 1561
QY 377 GTCAAGGCTTCAAGTGTCAAGGCAATGACCAACCAAGAAAGCTTAAAGCTGTGAG 436
Db 1562 GTCAAGGCTTCAAGTGTCAAGGCAATGACCAACCAAGAAAGCTTAAAGCTGTGAG 1621
QY 437 TTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGGAGTGGGCGCTTGAAGAGA 496
Db 1622 TTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGGAGTGGGCGCTTGAAGAGA 1681
QY 497 CCCAAGAGTATCTCAAGTGTGAGAGAAACCTGTGAGTTAACTGACGATATGCCAGA 556
Db 1682 CCCAAGAGTATCTCAAGTGTGAGAGAAACCTGTGAGTTAACTGACGATATGCCAGA 1741
QY 557 GATCCCCGAAATTCACCTCAAGAAACCTTTGGAACAATCACCAACCAATTCACATCAC 616
Db 1742 GATCCCCGAAATTCACCTCAAGAAACCTTTGGAACAATCACCAACCAATTCACATCAC 1801
QY 617 ACTCAAGATCAATGATCTTATTGGGCTTGGAAAAAGTCTG 663
Db 1802 ACTCAAGATCAATGATCTTATTGGGCTTGGAAAAAGGATG 1848

Search completed: March 25, 2004, 05:47:07
Job time : 8608 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 01:49:44 ; Search time 855 Seconds

(without alignments)
11010.534 Million cell updates/sec

Title: US-09-622-964-1_COPY_13900_16115

Perfect score: 2216
Sequence: 1 attctctgtctcttcacgc.....ctaatgagtttaataaac 2216

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

N_Geneseq_29Jan04: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2003as: *
8: geneseqn2003bs: *
9: geneseqn2003cs: *
10: geneseqn2004s: *

Pred. No. is the number of results predicted by chance (i.e. have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2215	100.0	16125	2	AAZ21226 Human CGI
2	2188.6	98.8	16650	5	ABAI4559 Human ner
3	2188.6	98.8	18530	5	ABAI4557 Human ner
4	2187	98.7	18537	5	ABAI4558 Human ner
5	1631.6	73.6	1717	6	ABLB9697 Human pol
6	643.8	29.1	1253	2	AAV99722 Human adu
7	642.2	29.0	1758	7	ABZ80972 Human bee
8	642.2	29.0	2229	2	AAZ21227 Human CGI
9	642.2	29.0	2429	2	AAZ21228 Human CGI
10	511.8	22.3	1238	8	ADAA4960 Human pol
11	494	22.3	7108	8	ABAI4556 Human ner
12	404.2	18.2	459	8	ACH39717 Human foe
13	399.4	18.0	424	2	AAV86976 Human foe
14	339.4	15.3	1326	7	ACA03889 cDNA upre
15	318	14.4	356	3	AAC20667 Human sec
16	164	7.4	1119	7	ABZ71972 Human fer
17	164	7.4	1198	6	ABK83853 Human cdn
18	164	7.4	1198	6	ABN95676 Human cdn
19	164	7.4	1234	5	AA667682 DNA encod
20	164	7.4	1279	5	AA685665 DNA encod
21	164	7.4	2183	5	AA691587 DNA encod
22	133.2	6.0	65042	7	AB855711 DNA encod
23	133.2	6.0	65042	8	ACA63029 Human gen

24	133.2	6.0	149480	6	ABL61947
25	133.2	6.0	149480	6	ABL68365
26	133.2	6.0	149480	6	ABL61948
27	127.6	5.8	322101	9	AAD58431
28	126.8	5.7	22900	4	AAK82210
29	126.2	5.7	22645	4	AAK89727
30	126.2	5.7	22645	4	AAI05495
31	126.2	5.7	22645	4	ABL04985
32	126.2	5.7	22645	4	ABL97878
33	126.2	5.7	22645	4	ABL98348
34	124.8	5.6	165199	6	ABK83460
35	124.8	5.6	177531	7	ACF62732
36	124.8	5.6	177531	7	ADB20847
37	124.8	5.6	177531	9	ADB87936
38	124.8	5.6	177531	9	ADB87936
39	124.8	5.6	177531	9	ADB96919
40	124.2	5.6	38997	6	AAD36069
41	124.2	5.6	700	4	AAH92912
42	124	5.6	96587	8	ADA02984
43	124	5.6	96587	9	ADB72722
44	124	5.6	96587	9	ADC85464
45	123.8	5.6	14874	4	AAK89726

ALIGNMENTS

RESULT 1	
AAZ21226	ID
AAZ21226	standard; DNA; 16125 BP.
XX	
AC	AAZ21226;
XX	
DT	22-NOV-1999 (first entry)
XX	
DE	Human CGICe genomic DNA sequence.
XX	
KW	CGICe; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
KW	age-related macular dystrophy; ss.
XX	
XX	Homo sapiens.
OS	
XX	
FH	Key
FT	intron
FT	Location/Qualifiers
FT	1..1513
FT	/*tag= a
FT	/number= 1
FT	1514..1570
FT	/*tag= b
FT	/number= 1
FT	1571..3278
FT	/*tag= c
FT	/number= 2
FT	3279..3466
FT	/*tag= d
FT	/number= 2
FT	3315..15787
FT	/*tag= e
FT	/product= "CGICe protein"
FT	/note= "Contains introns"
FT	3467..6616
FT	/*tag= f
FT	/number= 3
FT	6617..6711
FT	/*tag= g
FT	/number= 3
FT	6712..7227
FT	/*tag= h
FT	/number= 4
FT	7228..7461
FT	/*tag= i
FT	/number= 1
FT	7462..8353
FT	/*tag= j
FT	intron

FT	/number= 5
FT	8354. .8508
FT	/*tag= k
FT	/number= 5
FT	8509. .8896
FT	/*tag= l
FT	/number= 6
FT	8897. .8974
FT	/*tag= m
FT	/number= 6
FT	8975. .9676
FT	/*tag= n
FT	/number= 7
FT	9677. .10032
FT	/*tag= o
FT	/number= 7
FT	10033. .11212
FT	/*tag= p
FT	/number= 8
FT	11213. .11293
FT	/*tag= q
FT	/number= 8
FT	11294. .11556
FT	/*tag= r
FT	/number= 9
FT	11557. .11708
FT	/*tag= s
FT	/number= 9
FT	11709. .13918
FT	/*tag= t
FT	/number= 10
FT	13919. .14557
FT	/*tag= u
FT	/number= 10
FT	14558. .15768
FT	/*tag= v
FT	/number= 11
FT	15769. .16125
FT	/*tag= w
XX	/number= 11
PX	
PN	WO9943695-A1.
XX	
PD	
XX	02-SEP-1999.
PF	
PR	22-FEB-1999; 99WO-US003790.)
XX	
PR	25-FEB-1998; 98US-0075941P.
PR	18-DEC-1998; 98US-0112926P.
XX	
PA	(MERI) MERCK & CO INC.
XX	(UYUP-) UNIV UPPSALA.
PI	
PS	Petrukhin K, Caskey CT, Metzker M, Wadellius C;
DR	WPI: 1999-540560/45.
XX	P-PSeDf; AAI29953.
PT	
Human and mouse polynucleotides encoding CGICE polypeptides. Claim 2; Fig 1; 67pp; English.	
The present sequence represents the human CGICE gene, which when mutated is responsible for Best's macular dystrophy (BMD). Polynucleotides encoding CGICE are useful for diagnosing whether a patient carries a mutation in the CGICE gene. Normal and mutated CGICE proteins are useful for identifying activators and/or inhibitors of these proteins, in order to treat BMD. The CGICE gene offers a simpler and cheaper method of diagnosing BMD without the need for the presence of the patient. The gene may also be useful to discovering the genetic cause of age-related macular dystrophy	
Sequence 16125 BP; 398A A; 4175 G; 4215 G; 3726 T; 0 U; 21 Other;	

Query Match	Best Local Similarity	100.0%	Score 2215	DB 2	Length 16125	Matches 2215	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ATCTCTGTTTCTTTCAGCTTGAA	CAAAAGAGAGATGAGTTTCCAGCCCAATCAGAG	60						
Db	13900	ATCTCTGTTTCTTTCAGCTTGAA	CAAAAGAGAGATGAGTTTCCAGCCCAATCAGAG	13955						
QY	61	ACGAGAGAGTGTCAAGCTGGCAT	TATTTGGCCGCTTCTTAGGCTTCAGTCCATGATC	120						
Db	13960	ACGAGAGAGTGTCAAGCTGGCAT	TATTTGGCCGCTTCTTAGGCTTCAGTCCATGATC	14019						
QY	121	ACCATCTTCCAGGGCAAACTGA	AGACCAACTCTAGTGGCCCAAGAGGGAATCCCTTC	180						
Db	14020	ACCATCTTCCAGGGCAAACTGA	AGACCAACTCTAGTGGCCCAAGAGGGAATCCCTTC	14079						
QY	181	TCACAGAGGCGCTGGCCAAAA	CCCAAGAGGAGCCCAACAGAGCTTAGGGCCAGAG	240						
Db	14080	TCACAGAGGCGCTGGCCAAAA	CCCAAGAGGAGCCCAACAGAGCTTAGGGCCAGAG	14139						
QY	241	ACAAACAAGCCTGGAAGCTTAA	GCTGTGAGCGCTTCAAGTCTGAGCCCACTGATACGA	300						
Db	14140	ACAAACAAGCCTGGAAGCTTAA	GCTGTGAGCGCTTCAAGTCTGAGCCCACTGATACGA	14199						
QY	301	GAGCAGGCTATACAGATGAGCC	CAAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCC	360						
Db	14200	GAGCAGGCTATACAGATGAGCC	CAAGAGCGCCCTCAGCCCACTCCATGTTCTTCCCC	14259						
QY	361	TAGAACCATCAGCGCGCTCAA	AGGTTTCAAGTGCACAGGCTATAGACCAAAAGACAAA	420						
Db	14260	TAGAACCATCAGCGCGCTCAA	AGGTTTCAAGTGCACAGGCTATAGACCAAAAGACAAA	14319						
QY	421	GCTTAAAGACTGAGTCTGGGG	CCCAAGAAAGTTTGAATTGCTCTCAGAGCGATG	480						
Db	14320	GCTTAAAGACTGAGTCTGGGG	CCCAAGAAAGTTTGAATTGCTCTCAGAGCGATG	14379						
QY	481	GGGCGTTTGAATGAGGACCA	CCCAAGATCTTCAAGTGAAGAGAAACTGTGAGTTTAACT	540						
Db	14380	GGGCGTTTGAATGAGGACCA	CCCAAGATCTTCAAGTGAAGAGAAACTGTGAGTTTAACT	14439						
QY	541	TGACGATATGCGAGATGCCG	AAATCACTTCAAGAGAACTTTGGAACAATACCA	600						
Db	14440	TGACGATATGCGAGATGCCG	AAATCACTTCAAGAGAACTTTGGAACAATACCA	14499						
QY	601	CCAACATACACTACCTCAAA	AGATCACATGATCTTATTGGGCTTGGAAAAAGGT	660						
Db	14500	CCAACATACACTACCTCAAA	AGATCACATGATCTTATTGGGCTTGGAAAAAGGT	14559						
QY	661	CTGTCTCTCACTGAAACAG	GGGCACTGCAATTGGCCCAACCCAGCTTCCCTTG	720						
Db	14560	CTGTCTCTCACTGAAACAG	GGGCACTGCAATTGGCCCAACCCAGCTTCCCTTG	14619						
QY	721	CTCTAGGCTTACCCCTTCCA	CAATTTCTTAGGGTTCAATCACTGTCAGAGCACTGG	780						
Db	14620	CTCTAGGCTTACCCCTTCCA	CAATTTCTTAGGGTTCAATCACTGTCAGAGCACTGG	14679						
QY	781	ACCTTACGCCAGCACTGGCT	TGGGTTATCTTGGCCACTTACAGAGGATCTTAGGA	840						
Db	14680	ACCTTACGCCAGCACTGGCT	TGGGTTATCTTGGCCACTTACAGAGGATCTTAGGA	14739						
QY	841	AGTGTGGGAGCTTTTCTCA	CTTCACTTCACTGGGATATCACCCGGAAGACTTCTTGGAGCAG	900						
Db	14740	AGTGTGGGAGCTTTTCTCA	CTTCACTTCACTGGGATATCACCCGGAAGACTTCTTGGAGCAG	14799						
QY	901	GTGAAGAAATAGAGTTGTG	TGACCAAGATCTGTGAGAACTGGCCCAAGGCTGAC	960						
Db	14800	GTGAAGAAATAGAGTTGTG	TGACCAAGATCTGTGAGAACTGGCCCAAGGCTGAC	14859						
QY	961	AGGCACAGGCTTATGAGAG	ATGTTATCACTTGGCCCACTTACTTTTGAAGAGGTGG	1020						
Db	14860	AGGCACAGGCTTATGAGAG	ATGTTATCACTTGGCCCACTTACTTTTGAAGAGGTGG	14919						


```

PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231245P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 25-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236377P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244222P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.

```

```

PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249219P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249248P.
PR 17-NOV-2000; 2000US-0249249P.
PR 17-NOV-2000; 2000US-0249289P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 01-DEC-2000; 2000US-0251160P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251899P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2001US-0259678P.

```

(HUMA-) HUMAN GENOME SCT INC.

Rosen CA, Barash SC, Ruben SM;
WPI; 2001-541565/60.

Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
PT and metastases.

XX Disclosure; SEQ ID NO 6890; 1701pp + Sequence Listing; English.

XX

CC The invention relates to novel genes (ABAI1004-ABA21534) and proteins
CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC / (e) neurological diseases such as viral, bacterial, fungal and parasitic
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at [ftp.wipo.int/pub/publicated_pct_sequences](http://wipo.int/pub/publicated_pct_sequences)

XX

Sequence 16650 BP; 4168 A; 4271 C; 4308 G; 3903 T; 0 U; 0 Other;

Query Match 98.8%; Score 2188.6; DB 5; Length 16650;
Best Local Similarity 99.5%; Pred No. 0;
Matches 2205; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

QY 1 ATCTCTGTTTCTTTCACGCTGAACTCAAGAGAGATGAGTTCCAGGCCATCAGAGG 60
Db 11879 ATCTCTGTTTCTTTCACGCTGAACTCAAGAGAGATGAGTTCCAGGCCATCAGAGG 11938

QY 61 ACAGAGAGATGCTCAACGCTGATCATTCATTCGCTTCTTCTGAGCTGATCCATGATC 120
Db 11939 ACAGAGAGATGCTCAACGCTGATCATTCATTCGCTTCTTCTGAGCTGATCCATGATC 11998

QY 121 ACCATCTCTCCAGGGCAACTCAAGACCAACTACTGTGCGCCCAAGAGGGAATCCCTTC 180
Db 11999 ACCATCTCTCCAGGGCAACTCAAGACCAACTACTGTGCGCCCAAGAGGGAATCCCTTC 12058

181 TCACGAGGGCTGCCCCCAACAGAGGCAAGCAAGAGCTTGGGGCAGAGAG 240
12059 TCACGAGGGGCTGCCCCCAACAGAGGCAAGCAAGAGCTTGGGGCAGAGAG 12118
241 ACAACAGGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGCCCCACTGTATGAGA 300
12119 ACAACAGGGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGCCCCACTGTATGAGA 12178
301 GGGCAGGCTACTACAGTGGCCCCACAGAGGCCCCCAGCCCACTCCCAAGTGTCTTCCCCC 360
12179 GGGCAGGCTACTACAGTGGCCCCACAGAGGCCCCCAGCCCACTCCCAAGTGTCTTCCCCC 12238
361 TAGAACCATCAGGCGCTCAAGAGCTTCAAGTGTGAGGCAATGAGACCAACAAAGACAAA 420
12239 TAGAACCATCAGGCGCTCAAGAGCTTCAAGTGTGAGGCAATGAGACCAACAAAGACAAA 12298
421 GCTTAAAGACTGTGAGTCTTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATG 480
12299 GCTTAAAGACTGTGAGTCTTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATG 12358
481 GGGCCTTGAATGAGGACCCAGAGATCTCTCAAGTGAAGAGAGAAACCTGTGAAGTTTAAAC 540
12359 GGGCCTTGAATGAGGACCCAGAGATCTCTCAAGTGAAGAGAGAAACCTGTGAAGTTTAAAC 12418
541 TGAAGGATATGCGAGAGATCCCGAAATATCACTCAAGAAACCTTGGAAATCAACAA 600
12419 TGAAGGATATGCGAGAGATCCCGAAATATCACTCAAGAAACCTTGGAAATCAACAA 12478
601 CCAACATACACACTACACTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
12479 CCAACATACACACTACACTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 12538
661 CTGTCTCCCACTGGAACCAAGGGGCACTGCAATGCGCTGTGCGCCCACTCCCTGCTTCC 720
12539 CTGTCTCCCACTGGAACCAAGGGGCACTGCAATGCGCTGTGCGCCCACTCCCTGCTTCC 12598
721 CTGTGAGCTTACCTTCTCCCACTTCTCAAGGGGCTTCACTGCTGCTGCTGCTGCTGCTGCT 780
12599 CTGTGAGCTTACCTTCTCCCACTTCTCAAGGGGCTTCACTGCTGCTGCTGCTGCTGCTGCT 12658
781 ACCTAGCCCGCAGCAGCTGGCTTGGGGTATATATCTTGGCCACTTTCACAGAGATCTTAAAG 840
12659 ACCTAGCCCGCAGCAGCTGGCTTGGGGTATATATCTTGGCCACTTTCACAGAGATCTTAAAG 12718
841 AGTGTTCGGGACCTTTTCTCACTTCACTGCTGTATACCCCGGAAGACTTCTTGGAGACAG 900
12719 AGTGTTCGGGACCTTTTCTCACTTCACTGCTGTATACCCCGGAAGACTTCTTGGAGACAG 12778
901 GTGAAGGAAGATGAGTGTGCTGACAGGAATGCTGTGGAAGATGCTGCTGCTGCTGCTGCTGCT 960
12779 GTGAAGGAAGATGAGTGTGCTGACAGGAATGCTGTGGAAGATGCTGCTGCTGCTGCTGCTGCT 12838
961 AGGCGAGGCTTGAAGCTGAGAGATGTTATCACTGCGCCCACTTATCTTGAAGAGGGTGG 1020
12839 AGGCGAGGCTTGAAGCTGAGAGATGTTATCACTGCGCCCACTTATCTTGAAGAGGGTGG 12898
1021 CTGAACCCAAACATGAGGCTGAGTCACTGCTGATGACAGATGAACCTTCCCGCATTAAC 1080
12899 CTGAACCCAAACATGAGGCTGAGTCACTGCTGATGACAGATGAACCTTCCCGCATTAAC 12958
1081 TATTTAGGAGTATACCAAGCACTACAGAAAGGGTGGAGAACTGCTCACTGCTGAG 1140
12959 TATTTAGGAGTATACCAAGCACTACAGAAAGGGTGGAGAACTGCTCACTGCTGAG 13018
1141 AACTGTAGATGATGAGTGTGAGGGTGTCCAGGCGCTTATGATCTTCTCACTGCTGCTG 1200
13019 AACTGTAGATGATGAGTGTGAGGGTGTCCAGGCGCTTATGATCTTCTCACTGCTGCTG 13078
1201 GGAACCTCACAACAAATCTTCTGCTTCTTGGGGTGTGAGCCCAAGTGTGCAACAAATCA 1260
13079 GGAACCTCACAACAAATCTTCTGCTTCTTGGGGTGTGAGCCCAAGTGTGCAACAAATCA 13138
1261 GATATTTCTCTTATTCAGATTTCTGAGACATGTCACCAATATATTAACACCCCACTT 1320

13139 GATATTTCTCTTATTCAGATTTCTGAGACATTTCAACCAATATTAACACCCCACTT 13198
1321 CAGCCCAATCAGGTGGAGAGAGTGTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1380
13199 CAGCCCAATCAGGTGGAGAGAGTGTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 13258
1381 TTCAAGGGTGAAGACAGCAGCTATTATGATTTGAAGCTTAAAGGGCAACAAATTTCAAT 1440
13259 TTCAAGGGTGAAGACAGCAGCTATTATGATTTGAAGCTTAAAGGGCAACAAATTTCAAT 13317
1441 CTGTCTTGAAGCTTGAAGACAGAACTTGGCAACATCTGTGGCTGTTCAGCAAGAGATG 1500
13318 CTGTCTTGAAGCTTGAAGACAGAACTTGGCAACATCTGTGGCTGTTCAGCAAGAGATG 13377
1501 TTCAATTTAAAGATTTGTCTTGGGCTGGGTGTGAGGCAATGATTCACAGAGGTCA 1560
13378 TTCAATTTAAAGATTTGTCTTGGGCTGGGTGTGAGGCAATGATTCACAGAGGTCA 13437
1561 GGAAGTTGAGACCACTGGCCAAATGATGAAGAACCCATCTGACCAAAATTAACAA 1620
13438 GGAAGTTGAGACCACTGGCCAAATGATGAAGAACCCATCTGACCAAAATTAACAA 13497
1621 ATCAGTGGCGCTGT 1680
13498 ATCAGTGGCGCTGT 13557
1681 TTGAACCCAGAGAGT 1740
13558 TTGAACCCAGAGAGT 13617
1741 GGAAGTGAAGCTGTCTCAAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1800
13618 GGAAGTGAAGCTGTCTCAAAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 13677
1801 GCAACATTTTGTATTTGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1860
13678 GCAACATTTTGTATTTGAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 13737
1861 TCTCAGTGAAGATGAAGACATTTCTTCACTGCTTCTTCAATGAGGATGCTTCCGACGCA 1920
13738 TCTCAGTGAAGATGAAGACATTTCTTCACTGCTTCTTCAATGAGGATGCTTCCGACGCA 13797
1921 GGTCTCAGCTGT 1980
13798 GGTCTCAGCTGT 13857
1981 ACACTGAAGACGCTGTCTTCAACAGCCTGGAATCAATGTTAGTTAATGAATTAAT 2040
13858 ACACTGAAGACGCTGTCTTCAACAGCCTGGAATCAATGTTAGTTAATGAATTAATGAAT 13917
2041 CCCAGACTTCAAGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 2100
13918 CCCAGACTTCAAGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCT 13977
2101 CCAATGGAACATTTAAATCTCAAGCTGGAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 2160
13978 CCAATGGAACATTTAAATCTCAAGCTGGAATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 14037
2161 AATCAAGACAGCAGCAGCTTAAATGATTAAGTCCCAATTAATGAATTAATGAATTAATGAAT 2216
14038 AATCAAGACAGCAGCAGCTTAAATGATTAAGTCCCAATTAATGAATTAATGAATTAATGAAT 14093

RESULT 3
ABAI4557
ID ABAI4557 strand; DNA, 18530 BP.
XX
AC ABAI4557;
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polynucleotide SEQ ID NO 6886.

XX Human; nootropic; neuroprotective; cytosstatic; dermatological; virocidic;
KW immunosuppressive; antitumeflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antidiabetic; antianemic; antitubercic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antitumeflammatory;
KW antiallergic; antidiabetic; antituber; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
XX neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205513P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-021647P.
PR 11-JUL-2000; 2000US-0216880P.
PR 14-JUL-2000; 2000US-0217487P.
PR 14-JUL-2000; 2000US-0217496P.
PR 26-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-022547P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 30-AUG-2000; 2000US-0227009P.
PR 01-SEP-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 05-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0233400P.

PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0234999P.
PR 27-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241828P.
PR 01-NOV-2000; 2000US-0242221P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246529P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.

PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-541565/60.
XX
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system cancers
XX and metastases.
XX
XX Disclosure; SEQ ID NO 6888; 1701bp + Sequence Listing; English.
XX
XX The invention relates to novel genes (AB11004-AB21534) and proteins
CC (AB14678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPD at fcp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 18530 BP; 4512 A; 4799 C; 4984 G; 4235 T; 0 U; 0 Other;
XX
XX Query Match 98.8%; Score 2188.6; DB 5; Length 18530;
XX Best Local Similarity 99.5%; Pred. No. 0;
XX Matches 2205; Conservative 0; Mismatches 10; Indels 1; Gaps 1;
XX
QY 1 ATCTCTGTTCTTTTCCAGCTGAACAAGAGATGAGTTCCAGCCCAATCAGAGG 60
DB 11879 ATCTCTGTTCTTTTCCAGCTGAACAAGAGATGAGTTCCAGCCCAATCAGAGG 11938
QY 61 ACGAGAGATGCTCAAGCTGCGATCATTTGGCGCTTCTTAGGCTGTAGTCCATGATC 120
DB 11939 ACGAGAGATGCTCAAGCTGCGATCATTTGGCGCTTCTTAGGCTGTAGTCCATGATC 11998
QY 121 ACCATCTCCAGGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTC 180
DB 11999 ACCATCTCCAGGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGAATCCCTTC 12058
QY 181 TCCACAGAGGCTGCCCCAAAACCAAGGACCAAGAACTTTAGGGGCGAGAG 240
DB 12059 TCCACAGAGGCTGCCCCAAAACCAAGGACCAAGAACTTTAGGGGCGAGAG 12118
QY 241 ACAACAAGGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAG 300
DB 12119 ACAACAAGGCTGGAAGCTTAAGGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAG 12178
QY 301 GGCAGGCTACTACAGTGGCCCAAGAGCGCCCTCAGGCCCACTCCATTTTCCCCC 360
DB 12179 GGCAGGCTACTACAGTGGCCCAAGAGCGCCCTCAGGCCCACTCCATTTTCCCCC 12238
QY 361 TAGAACCATATGAGCGGCTCAAGAGTTTCAAGTGTACAGGCAATGACCAAAAGACAAA 420
DB 12239 TAGAACCATATGAGCGGCTCAAGAGTTTCAAGTGTACAGGCAATGACCAAAAGACAAA 12298
QY 421 GCTTAAGACTGTAGTTCGTGGGCGCAAGAAAGTTTGAATTGATTCAGAGAGGATG 480

DB 12299 GCTTAAGACTGTAGTTCGTGGGCGCAAGAAAGTTTGAATTGATTCAGAGAGGATG 12358
QY 481 GGCCTTTGATGAGCAACCAAGAACTTCTCAAGTGAAGAGAAACTGTGAGTTTAAAC 540
DB 12359 GGCCTTTGATGAGCAACCAAGAACTTCTCAAGTGAAGAGAAACTGTGAGTTTAAAC 12418
QY 541 TGAAGGATATGCAAGATCCCGGAAATCACTCAAGAACTTTGGAACTATCAACCA 600
DB 12419 TGAAGGATATGCAAGATCCCGGAAATCACTCAAGAACTTTGGAACTATCAACCA 12478
QY 601 CCAACATACACACTACACTCAAAAGATCACAATGATCTTATGGCCCTTGGAAACAGT 660
DB 12479 CCAACATACACACTACACTCAAAAGATCACAATGATCTTATGGCCCTTGGAAACAGT 12538
QY 661 CTGTCTCCACCTGAACCAAGGAGCTGCAATGCTGTCGCCCAACCCAGCTTCCCTTG 720
DB 12539 CTGTCTCCACCTGAACCAAGGAGCTGCAATGCTGTCGCCCAACCCAGCTTCCCTTG 12598
QY 721 CTGTAGGCTTACCTTCTCAAAATTTCTAGGGTTCACTGACAGAGCACTGG 780
DB 12599 CTGTAGGCTTACCTTCTCAAAATTTCTAGGGTTCACTGACAGAGCACTGG 12658
QY 781 ACCTAGCCCAAGCACTGCTGGGGTATATCTTGGCCCACTTCAAGGAGCTTGAAGA 840
DB 12659 ACCTAGCCCAAGCACTGCTGGGGTATATCTTGGCCCACTTCAAGGAGCTTGAAGA 12718
QY 841 AGTGTGGGACCTTTTCTCACTTCAACCTGATATCAACCGGAAGACTTTGGAGACAG 900
DB 12719 AGTGTGGGACCTTTTCTCACTTCAACCTGATATCAACCGGAAGACTTTGGAGACAG 12778
QY 901 GTGAAGAAATGAGGTTGCTGACAGAAATGCTGTGAGAACTGGCCCAAGGCTGAC 960
DB 12779 GTGAAGAAATGAGGTTGCTGACAGAAATGCTGTGAGAACTGGCCCAAGGCTGAC 12838
QY 961 AGGCGAGCTTACCTGAGAGATGTTATCACTGGCCCACTTACCTTGAAGAAAGGTGG 1020
DB 12839 AGGCGAGCTTACCTGAGAGATGTTATCACTGGCCCACTTACCTTGAAGAAAGGTGG 12898
QY 1021 CTGACCCAAAACATGAGGTGCGATGACGTGATGACAGATGACACTTCCCCATTAAC 1080
DB 12899 CTGACCCAAAACATGAGGTGCGATGACGTGATGACAGATGACACTTCCCCATTAAC 12958
QY 1081 TATTTAGGTAAGTACCAAGCACTACAGAAAGGAGGAGAACTGCCCACTCTCAG 1140
DB 12959 TATTTAGGTAAGTACCAAGCACTACAGAAAGGAGGAGAACTGCCCACTCTCAG 13018
QY 1141 AACTGTAGATGATGAGGTTGAGGGTGTCAAGGCGCTTAAAGTCACTTCTCACTGCTG 1200
DB 13019 AACTGTAGATGATGAGGTTGAGGGTGTCAAGGCGCTTAAAGTCACTTCTCACTGCTG 13078
QY 1201 GGAACCTCAACAAATACTTCTTGTCTTCTTGGGGTCAAGCCCAAGCTGTCAAAATCA 1260
DB 13079 GGAACCTCAACAAATACTTCTTGTCTTCTTGGGGTCAAGCCCAAGCTGTCAAAATCA 13138
QY 1261 GATATTTCCCTTATTTCCAGATTTCTGAGCACTGTCAACCAATTTATTAACACCCACTT 1320
DB 13139 GATATTTCCCTTATTTCCAGATTTCTGAGCACTGTTCACCAATTTATTAACACCCACTT 13198
QY 1321 CAGCCCAATCAAGTGGAGAAAGTAACTTCTTCTTGTGATTTCTCAAGAGTAACT 1380
DB 13199 CAGCCCAATCAAGTGGAGAAAGTAACTTCTTCTTGTGATTTCTCAAGAGTAACT 13258
QY 1381 TTCACGGGTCAAGAACGAGCTATTATGATGAAAACCTTAAAGGGCAACAAATTCANT 1440
DB 13259 TTCAC-GGTCAAGAACGAGCTATTATGATGAAAACCTTAAAGGGCAACAAATTCAGT 13317
QY 1441 CTTCCTTTAGGCTAAGACAGAACTTGGCAACATCTGTGGCTTTTCAAGCAAGAGT 1500
DB 13318 CTTCCTTTAGGCTAAGACAGAACTTGGCAACATCTGTGGCTTTTCAAGCAAGAGT 13377
QY 1501 TTCAATTTAAGATCTTGTCTTGGGCTGGGTGTGAGAGCAAGTAACTACAGAGATCA 1560

PR	31-JAN-2000	2000US-01790655
PR	04-FEB-2000	2000US-0160628P
PR	24-FEB-2000	2000US-0184664P
PR	02-MAR-2000	2000US-0186350P
PR	16-MAR-2000	2000US-01898743P
PR	17-MAR-2000	2000US-0190076P
PR	18-APR-2000	2000US-0198123P
PR	19-MAY-2000	2000US-02005515P
PR	07-JUN-2000	2000US-02005457P
PR	28-JUN-2000	2000US-0214986P
PR	30-JUN-2000	2000US-0214986P
PR	07-JUL-2000	2000US-0216547P
PR	11-JUL-2000	2000US-02174887P
PR	14-JUL-2000	2000US-0217496P
PR	26-JUL-2000	2000US-0218290P
PR	26-JUL-2000	2000US-0220963P
PR	14-AUG-2000	2000US-0220964P
PR	14-AUG-2000	2000US-0224518P
PR	14-AUG-2000	2000US-0224519P
PR	14-AUG-2000	2000US-02252513P
PR	14-AUG-2000	2000US-02252514P
PR	14-AUG-2000	2000US-0225566P
PR	14-AUG-2000	2000US-0225667P
PR	14-AUG-2000	2000US-0225706P
PR	14-AUG-2000	2000US-0225747P
PR	14-AUG-2000	2000US-0225547P
PR	14-AUG-2000	2000US-0225556P
PR	18-AUG-2000	2000US-0225759P
PR	22-AUG-2000	2000US-0226881P
PR	22-AUG-2000	2000US-0226868P
PR	23-AUG-2000	2000US-0227182P
PR	23-AUG-2000	2000US-0227099P
PR	30-AUG-2000	2000US-0228924P
PR	01-SEP-2000	2000US-0228287P
PR	01-SEP-2000	2000US-0229343P
PR	01-SEP-2000	2000US-0229345P
PR	05-SEP-2000	2000US-0229345P
PR	05-SEP-2000	2000US-0229509P
PR	06-SEP-2000	2000US-0230431P
PR	06-SEP-2000	2000US-0230438P
PR	08-SEP-2000	2000US-0231242P
PR	08-SEP-2000	2000US-0231243P
PR	08-SEP-2000	2000US-0231244P
PR	08-SEP-2000	2000US-0231413P
PR	08-SEP-2000	2000US-0231414P
PR	08-SEP-2000	2000US-0232080P
PR	12-SEP-2000	2000US-0232081P
PR	14-SEP-2000	2000US-0232338P
PR	14-SEP-2000	2000US-0232339P
PR	14-SEP-2000	2000US-0232400P
PR	14-SEP-2000	2000US-0232401P
PR	14-SEP-2000	2000US-0233063P
PR	14-SEP-2000	2000US-0233064P
PR	21-SEP-2000	2000US-0234223P
PR	21-SEP-2000	2000US-0234223P
PR	25-SEP-2000	2000US-0234377P
PR	25-SEP-2000	2000US-0234998P
PR	26-SEP-2000	2000US-0235484P
PR	27-SEP-2000	2000US-0235834P
PR	27-SEP-2000	2000US-0235836P
PR	29-SEP-2000	2000US-0236327P
PR	29-SEP-2000	2000US-0236367P
PR	29-SEP-2000	2000US-0236368P
PR	29-SEP-2000	2000US-0236569P
PR	29-SEP-2000	2000US-0236670P
PR	02-OCT-2000	2000US-0236602P

PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239355P.
 PR 13-OCT-2000; 2000US-0239356P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 01-NOV-2000; 2000US-0244231P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 17-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249301P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 01-DEC-2000; 2000US-0251160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251899P.
 PR 08-DEC-2000; 2000US-0251900P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX
 DR MPI; 2001-541565/60.
 XX
 PT Nucleic acids encoding 3224 human nervous system antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating nervous system cancers
 and metastases.

XX
 PS Disclosure; SEQ ID NO 6889; 1701bp + Sequence Listing; English.
 XX
 CC The invention relates to novel genes (ABAI1004-ABA21534) and proteins
 CC (ABAI4678-ABAI8001) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (anti)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 18537 BP; 4519 A; 4800 C; 4983 G; 4235 T; 0 U; 0 Other;
 Query Match 98.7%; Score 2187; DB 5; Length 18537;
 Best Local Similarity 99.5%; Pred. No. 0;
 Matches 2204; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
 QY 1 ATCTCTGTTCTTTTCCAGCTGTAACAAAGAGATGAGTTCCAGCCCAATCAGAGG 60
 DB 11885 ATCTCTGTTCTTTTCCAGCTGTAACAAAGAGATGAGTTCCAGCCCAATCAGAGG 11944
 QY 61 ACAGAGAGATGTCACGCTGAGATGATGCGGCTTCCATGAGCCGATCCCATATC 120
 DB 11945 ACAGAGAGATGTCACGCTGAGATGATGCGGCTTCCATGAGCCGATCCCATATC 12004
 QY 121 ACCATCTCCCAAGGGAATCTCAAGACCAACTAGTGTGCGCAAGAGGAAATCCCTTC 180
 DB 12005 ACCATCTCCCAAGGGAATCTCAAGACCAACTAGTGTGCGCAAGAGGAAATCCCTTC 12064
 QY 181 TCACAGAGGCGCTGCCCAAAACCAAGGACGACCAAGCAAGACGTTAGGGGCCAGAG 240
 DB 12065 TCACAGAGGCGCTGCCCAAAACCAAGGACGACCAAGCAAGACGTTAGGGGCCAGAG 12124
 QY 241 ACAACAAGGCGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGA 300
 DB 12125 ACAACAAGGCGCTGGAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGA 12184
 QY 301 GGCAGGCTACTACAGTGCACCAAGAGCGCCCTCAGGCCCATCCCATGTTCTTCCGCC 360
 DB 12185 GGCAGGCTACTACAGTGCACCAAGAGCGCCCTCAGGCCCATCCCATGTTCTTCCGCC 12244
 QY 361 TAGAACCATCAGGCGCGTCAAGCTTCAAGTGTCAAGGATAGACCAAGAACAA 420
 DB 12245 TAGAACCATCAGGCGCGTCAAGCTTCAAGTGTCAAGGATAGACCAAGAACAA 12304
 QY 421 GCTTAAAGACTGTGAATTTCTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGCATG 480
 DB 12305 GCTTAAAGACTGTGAATTTCTGGGCGCAAGAAAGTTTGAATGCTCTCAGAGAGCATG 12364
 QY 481 GGGCTTGAATGAGCAACCAAGATATCTCAAGTGAAGAGAAACGTTGAGATTTAAC 540
 DB 12365 GGGCTTGAATGAGCAACCAAGATATCTCAAGTGAAGAGAAACGTTGAGATTTAAC 12424
 QY 541 TGACGATATGCCAGATGCCGGAATTCACCTCAAGAACTTTGGAACAATCAACAA 600
 DB 12425 TGACGATATGCCAGATGCCGGAATTCACCTCAAGAACTTTGGAACAATCAACAA 12484
 QY 601 CCAACATACACTACACTCAAGATCAATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 12485 CCAACATACACTACACTCAAGATCAATGATGATGATGATGATGATGATGATGATGATGAT 12544
 QY 661 CTGTCTCCACCTGAACCAAGGCGCATGATGCGCTGTGCGCCACCCAGCTTCCCTTG 720

D 12545 CTGTCCTCCACCTGAAACAGAGGCACTGATTCCTGTGCCCAACCCCAAGCTTCCTTG 12604
Q 721 CTCTGAGCTTACCTCTCTCCCAAAATTCCTAGGGTCCATCATCTGACGACGACACCTGG 780
D 12605 CTCTGAGCTTACCTCTCTCCCAAAATTCCTAGGGTCCATCATCTGACGACGACACCTGG 12664
Q 781 ACCTAGCCCAACGACCTGGCTGGGGTATTAATCTTGGCCACTTCAACAGGATCCCTAGGA 840
D 12665 ACCTAGCCCAACGACCTGGCTGGGGTATTAATCTTGGCCACTTCAACAGGATCCCTAGGA 12724
Q 841 AGTGTTCGGGACCTTTTCTCACTTCACTCCCTGTATCAACCGGAAAGACTTCTTGGGACAG 900
D 12725 AGTGTTCGGGACCTTTTCTCACTTCACTCCCTGTATCAACCGGAAAGACTTCTTGGGACAG 12784
Q 901 GTGAAGAAAGATGAGAGTGTGTCTGACCAAGATGCTGTGGGAACTGCCCCAGGGCTGAC 960
D 12785 GTGAAGAAAGATGAGAGTGTGTCTGACCAAGATGCTGTGGGAACTGCCCCAGGGCTGAC 12844
Q 961 AGGCCAGCTTACCTGAGCAGATGTATCACTGCCCCCACTTACTTGTGAGCAAGGGTGG 1020
D 12845 AGGCCAGCTTACCTGAGCAGATGTATCACTGCCCCCACTTACTTGTGAGCAAGGGTGG 12904
Q 1021 CTGACCCAAAACATGAGGGTGGGAGTGAATGACAGATGACACTTCCCAATAC 1080
D 12905 CTGACCCAAAACATGAGGGTGGGAGTGAATGACAGATGACACTTCCCAATAC 12964
Q 1081 TATTTAGGTAATCCCAAGCACTATCAAGAAAGGGTGGCAGGAACTGCTCTCCTTAGG 1140
D 12965 TATTTAGGTAATCCCAAGCACTATCAAGAAAGGGTGGCAGGAACTGCTCTCCTTAGG 13024
Q 1141 AACTGTAGATGAGTGAAGTGTGAGGGTTCAGAGCCCTTAAAGTCAATTTTCTCACTGCTG 1200
D 13025 AACTGTAGATGAGTGAAGTGTGAGGGTTCAGAGCCCTTAAAGTCAATTTTCTCACTGCTG 13084
Q 1201 GGAACCTCACAAAATTAATCTTCTTCTTGGGGTCAAGCCAAAGCTGTCAAAAATCA 1280
D 13085 GGAACCTCACAAAATTAATCTTCTTCTTGGGGTCAAGCCAAAGCTGTCAAAAATCA 13144
Q 1261 GATATTTCCCTTATTTCCAGATTTCTGTGACACTGTCAACCAATTAATTAACCCCACTT 1320
D 13145 GATATTTCCCTTATTTCCAGATTTCTGTGACACTGTCAACCAATTAATTAACCCCACTT 13204
Q 1321 CAGCCCCAATCACTGGGAGAGAGTGAATCTCCCTTTTGTGATTTCTCAAGCAAGTTACT 1380
D 13205 CAGCCCCAATCACTGGGAGAGAGTGAATCTCCCTTTTGTGATTTCTCAAGCAAGTTACT 13264
Q 1381 TTCAAGGGTCAAGAACGACGACTATTAATGATTAATTAATTAATTAATTAATTAATTAAT 1440
D 13265 TTCAAGGGTCAAGAACGACGACTATTAATGATTAATTAATTAATTAATTAATTAATTAAT 1440
Q 1441 CTGCTTCTTAAAGCTAATGAGAACTTGGCAAACTGTGGGCTGTGTGACGAAAGAGATG 1500
D 13324 CTGCTTCTTAAAGCTAATGAGAACTTGGCAAACTGTGGGCTGTGTGACGAAAGAGATG 13383
Q 1501 TTCAATTTTAAAGATCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATCAAGAGGTCA 1560
D 13384 TTCAATTTTAAAGATCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATCAAGAGGTCA 13443
Q 1561 GAGCTTGAAGCAACTGTGGCAAACTGATGAATGAATCCCATCTCTCAACCAAAAAATTAACA 1620
D 13444 GAGCTTGAAGCAACTGTGGCAAACTGATGAATGAATCCCATCTCTCAACCAAAAAATTAACA 13503
Q 1621 ATCAGCTGGCCGTGCTGTGTGTCTGTAGTCCCAACGACGAGGTTGAGGGGAGAAATTCG 1680
D 13504 ATCAGCTGGCCGTGCTGTGTGTCTGTAGTCCCAACGACGAGGTTGAGGGGAGAAATTCG 13563
Q 1681 TTGAACCCGAGAGGT 1740
D 13564 TTGAACCCGAGAGGT 13623
Q 1741 GAGGTGAGACTGTCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
D 13624 GAGGTGAGACTGTCTCAAAAAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 13683

Q 1801 GCACATTTTGGATATTTGAATGAAGGATCCTTCATCTTATGCTGTAAATCTTCAAT 1860
D 13684 GCACATTTTGGATATTTGAATGAAGGATCCTTCATCTTATGCTGTAAATCTTCAAT 13743
Q 1861 TCTCCTTGGGATGAAGCAATTCCTTACCTGCTTCTTAATGGGATGCTTGGCAGCCA 1920
D 13744 TCTCCTTGGGATGAAGCAATTCCTTACCTGCTTCTTAATGGGATGCTTGGCAGCCA 13803
Q 1921 GGTCTCACTGTGTGTGACACGACGAGACACTGAATCCAGTCAACAGCCATACAGCTGTCC 1980
D 13804 GGTCTCACTGTGTGTGACACGACGAGACACTGAATCCAGTCAACAGCCATACAGCTGTCC 13863
Q 1981 ACACTGAAGAACTGTCTTCAACACAGCTGAATCAAAATGTTAGTTAATGAATTAAT 2040
D 13864 ACACTGAAGAACTGTCTTCAACACAGCTGAATCAAAATGTTAGTTAATGAATTAAT 13923
Q 2041 CCCAGACTTCACTGAGCTTAAATGCTTTTATTTCAATTAATAAATCTGAAAGCTGAA 2100
D 13924 CCCAGACTTCACTGAGCTTAAATGCTTTTATTTATTAATAAATCTGAAAGCTGAA 13983
Q 2101 CCATTTGGAACATTTTACTCACTGAGCTGTGATTCAGAGTGGGAAACCTTATGCTG 2160
D 13984 CCATTTGGAACATTTTACTCACTGAGCTGTGATTCAGAGTGGGAAACCTTATGCTG 14043
Q 2161 AATCCAAAGACACCAACCTTACTGCTTACTGCTTAACTAATGAGTTAATTAATTAAT 2216
D 14044 AATCCAAAGACACCAACCTTACTGCTTACTGCTTAACTAATGAGTTAATTAATTAAT 14099

RESULT 5
ABL9697
ID ABL9697 standard; cDNA; 1717 BP.
AC ABL9697;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polynucleotide SEQ ID NO 259.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatocytotoxic; antidiabetic; antiinflammatory; antileuk;
KW vulnereary; anticonvulsant; antibacterial; antitubercular; antiparasitic;
KW cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein; gene; ss.
OS Homo sapiens.
XX
PN W0200190304-A2.
XX
PD 29-NOV-2001.
XX
PF 18-MAY-2001; 2001MO-US016450.
XX
PR 19-MAY-2000; 2000US-020551SP.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-122018/16.
XX
DR P-PSDB; ABB89288.
XX
PT Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX disorders.
PS
XX Claim 4; SEQ ID NO 259; 2081bp + Sequence Listing; English.
CC
CC The invention relates to novel genes (ABL9449-ABL9085) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are

OS Homo sapiens.
XX

/*tag= a

PD 17-DEC-1998
XX

05-JUN-1998; 98US-00092722.
FK
XX

PI	Agostino MJ, Howes SH, Fechtel K,	Racie LA, Treacy M, Spaulding V,
XX		

PI New polynucleotides encoding secreted human proteins - derived from human
PPT foetal brain, adult testes, foetal kidney, adult thyroid or adult retina
PPT foetal brain, adult testes, foetal kidney, adult thyroid or adult retina
PPT foetal brain, adult testes, foetal kidney, adult thyroid or adult retina
CDNA libraries.

open reading frame for a 261-amino acid protein designated AAV95345. The clone was isolated from a human adult retina cDNA library using methods which are selective for cDNAs encoding secreted proteins. The cDNA identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoding protein. Database searches indicate some sequence similarity to known sequences. The invention provides cDNA clones (see AAV9271-33) from human adult thyroid, adult retina, adult testis, foetal kidney and foetal brain that encode novel secreted proteins (see AAV9534-45). Each clone is individually available from deposit clone ATCC 98451 (see also AAV9734-43). The isolated polynucleotides (PNTs) and proteins are predicted to have activities which would make them suitable for treating, preventing or ameliorating medical conditions in humans and animals, although no supporting cell is given. Suggested activities include nutritional, cytokine, cell proliferation or differentiation, immune stimulating (e.g. as vaccines) or immune suppressing, haematopoietic regulating (e.g. growth), activin/inhibin, chemotactic/chemokinetic, haemostatic, tissue thrombolytic, receptor/ligand, antiinflammatory, cathexin/tumour invasion suppressor, and tumour inhibition activities. The PNTs are also stated to be useful for gene therapy.

Match	Local Similarity	Pred. No.	Conservative	Mismatches	Length
645	99.7%	1.6e-180	0	0	1263

	mismatches	0;	Gaps	0;
17	Indels	2;		

D5
225 CAGCCTGAACAAGAGGAGATGGAGTCCAGGCCAATACCGCACCA

[illegible][illegible]

...CAAGAGGAATCCCTTCTCCACGAGGCTGCC 404

405 CAGGAGACACACAGCCTGGAA 256

257 GCTTA AGCCTGCTGAA 464

[illegible]

QY 317 TGCCCAAGACGCCCTTGAACCACTCACAATGCTGGCGTTTC

D_b
525 TGCCCAAGAGGCCCTCAGCCCCACTCCCATTGTTCTTCCCCCTAATCGGTTCCTTG

377 GTC AAGCTT CACAGTGT CACAGGCATAGACACCAAGACA AAGGTTTAAAGCTCTGTAAG

585 GTCAAAGCTTCACAGTGTCAAGGCATAGACACCAAGACAAAGCTTAAGACTGTAC 644

43 / TCTGGGGCCAGAAAAGTTTGAATTGCTCTCAGAGAGCGATGGGCCCTTGATGGACA 495

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

CTGTGAGCTTAACCTGACGGATATGCCAGA 556

...GAGGTTAACTGACCGATATGCCAGA 764

765 CATTCCCGCTTTTTCATGAGACAAI CACC AACCA CATACACTAC 616

617 ACTCAAGATGATTGTTGGTCTTGTTCTT
.....CACCACCAACAACACTAC 824

825 ACTCAAGATCACATGGATCTCTTAAATTCCTGTTA

BZ80972

ABZ80677.

15-OCT-2003 (Friday)

Human bestrophin (vital) form new...

Bestrophin; vitelliform macular dystrophy

antihypertensive; antidiabetic; immunomodulator;

eating disorder; cachexia; diabetes mellitus; hypertension; body-weight regulation;

sleep apnea; chromosome 11, osteoarthritis; cancer;

hollis sapiens.

no	Location/Qualifiers
CDS	1 1758

```

/ - category = d
/product = "VMP2"

```

WO2003030922-A2.

17-APR-2003.

09-OCT-2002; 2002WO-EP011321.

09-UCI-2001; 2001EP-00124059.

REVIEW / DEVELOPMENTS IN DEVELOPMENTAL BIOLOGICAL RESEARCH.

Proemer G, Fritsch R, Eulenberg K, Ciossek T;

P-PSDB; ABR58055.

PT New pharmaceutical compositions comprising a Bestrophin gene, polypeptide
PT or nucleic acid, for treating, alleviating and/or preventing metabolic
PT diseases, e.g. obesity, cachexia, diabetes mellitus, hypertension, or
PT galactosemia.

XX Claim 2; Fig 4a; 85bp; English.

XX This sequence represents the coding region for a member of the human
XX bestrophin gene family designated vitelliform macular dystrophin 2
XX (VMD2). The dystrophin gene family are involved in energy homeostasis and
XX metabolism of triglycerides. The sequence can be used for the manufacture
XX of an agent for detecting and/or verifying, for treating, alleviating
XX and/or preventing disorders including metabolic diseases such as obesity
XX and other body-weight regulation and related disorders such as eating
XX disorder, cachexia, diabetes mellitus, hypertension, coronary heart
XX disease, hypercholesterolemia, osteoarthritis, galactosemia, cancers of the
XX reproductive organs, and sleep apnea. The gene is found on human
XX chromosome 11

XX Sequence 1758 BP; 421 A; 515 C; 448 G; 374 T; 0 U; 0 Other;

Query Match 29.0%; Score 642.2; DB 7; Length 1758;
Best Local Similarity 99.5%; Pred. No. 5.8e-180;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 17 CAGCCTGAACAAGAGAGATGAGTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 76
DB 1098 CAGCCTGAACAAGAGAGATGAGTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 1157
QY 77 CGGTGGCATCATTTGGCGCTTCTCTAGGCTGAGTCCATGATCAGCATCTCCAGAGGC 136
DB 1158 CGGTGGCATCATTTGGCGCTTCTCTAGGCTGAGTCCATGATCAGCATCTCCAGAGGC 1217
QY 137 AAATCTAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCGCTGCC 196
DB 1218 AAATCTAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCGCTGCC 1277
QY 197 CAAAAACCAAGGAGGAGCAAG 256
DB 1278 CAAAAACCAAGGAGGAGCAAG 1337
QY 257 GCTTAAGGCTGTGAGAGCCTTCAAGTCTGCCCCAAGTATCAGAGAGGAGGCTACTACAG 316
DB 1338 GCTTAAGGCTGTGAGAGCCTTCAAGTCTGCCCCAAGTATCAGAGAGGAGGCTACTACAG 1397
QY 317 TGGCCCAAGAGAGGCGCCCTCAAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCC 376
DB 1398 TGGCCCAAGAGAGGCGCCCTCAAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCC 1457
QY 377 GTCAAGCTTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 436
DB 1458 GTCAAGCTTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAGAGTCAAG 1517
QY 437 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGGAGTGGGCTTGTATGAGCA 496
DB 1518 TTCTGGGGCCCAAGAAAGTTTGAATTGCTCTCAGAGAGGAGTGGGCTTGTATGAGCA 1577
QY 497 CCCAGAAGTATCTCAAGTGAAGAGAAACTGTGAGATTTAACCTGAACGATATGCGAGA 556
DB 1578 CCCAGAAGTATCTCAAGTGAAGAGAAACTGTGAGATTTAACCTGAACGATATGCGAGA 1637
QY 557 GATCCCGGAAATCACTCAAGAAAGTCTTGAAGCAATCAAGCAAGCAAGCAAGCAAGCAAGCA 616
DB 1638 GATCCCGGAAATCACTCAAGAAAGTCTTGAAGCAATCAAGCAAGCAAGCAAGCAAGCAAGCA 1697
QY 617 ACTCAAGATCAGATGATCTTATTTGGGCTTGAAGAAAGAGGCTCTG 663
DB 1698 ACTCAAGATCAGATGATCTTATTTGGGCTTGAAGAAAGAGGATG 1744

RESULT 8
AA221227
ID AA221227 standard; cDNA; 2229 BP.

XX AC AA221227;
XX DT 22-NOV-1999 (first entry)
XX DE Human CGICE short form cDNA sequence.
XX KW CGICE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
XX age-related macular dystrophy; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX FT CDS 105..1862

XX FT /tag= a
XX /product= "CGICE long form protein"
XX /tranel_except= (pos:465..467,aa:Ser)

XX MO9943695-A1.

XX PN 02-SEP-1999.

XX PF 22-FEB-1999; 99MO-US003790.

XX PR 25-FEB-1998; 98US-0075941P.

XX PR 18-DEC-1998; 98US-0112926P.

XX (MERI) MERCK & CO INC.

XX (UYUP-) UNIV UPPSALA.

XX PI Petrunkhin K, Caskey CT, Metzker M, Wadelius C;

XX WPI; 1999-540560/45.

XX DR P-PDB; AAY2953.

XX PT Human and mouse polynucleotides encoding CGICE polypeptides.

XX PS Claim 2; Fig 2; 67bp; English.

XX CC The present sequence represents the human CGICE cDNA sequence, which when

XX CC mutated is responsible for Best's macular dystrophy (BMD).

XX CC Polynucleotides encoding CGICE are useful for diagnosing whether a

XX CC patient carries a mutation in the CGICE gene. Normal and mutated CGICE

XX CC proteins are useful for identifying activators and/or inhibitors of these

XX CC proteins, in order to treat BMD. The CGICE gene offers a simpler and

XX CC cheaper method of diagnosing BMD without the need for the presence of the

XX CC patient. The gene may also be useful to discovering the genetic cause of

XX CC age-related macular dystrophy

XX SQ Sequence 2229 BP; 575 A; 646 C; 532 G; 476 T; 0 U; 0 Other;

Query Match 29.0%; Score 642.2; DB 2; Length 2229;
Best Local Similarity 99.5%; Pred. No. 6.7e-180;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 17 CAGCCTGAACAAGAGAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTCA 76
DB 1202 CAGCCTGAACAAGAGAGATGAGTCCAGCCCAATCAGAGAGAGAGAGATGCTCA 1261
QY 77 CGGTGGCATCATTTGGCGCTTCTCTAGGCTGAGTCCATGATCAGCATCTCCAGAGGC 136
DB 1262 CGGTGGCATCATTTGGCGCTTCTCTAGGCTGAGTCCATGATCAGCATCTCCAGAGGC 1321
QY 137 AAATCTAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCGCTGCC 196
DB 1322 AAATCTAAGAGCAAACTACTGTGGCCCAAGAGGAAATCCCTTCTCCAGAGGCGCTGCC 1381
QY 197 CAAAAACCAAGGAGGAGCAAG 256
DB 1382 CAAAAACCAAGGAGGAGCAAG 1441
QY 257 GCTTAAGGCTGTGAGAGCCTTCAAGTCTGCCCCAAGTATCAGAGGAGGCTACTACAG 316

Db 1442 GCTTAAGGCTGTGAGCGCCTTCAAGTCTGCGCCCACTGATATCAGAGGCCAGCTTACAG 1501
Qy 317 TGGCCACAGAGCGCCCTTCAAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCGC 376
Db 1502 TGCCCAAGAGAGCGCCCTTCAAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCGC 1561
Qy 377 GTCAAGGCTTCAAGTGTCAAGGATATGACACCAAGAACAAAGCTTAAAGCTGTAG 436
Db 1562 GTCAAGGCTTCAAGTGTCAAGGATATGACACCAAGAACAAAGCTTAAAGCTGTAG 1621
Qy 437 TTTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGGATGAGGCTTATGAGCA 496
Db 1622 TTTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGGATGAGGCTTATGAGCA 1681
Qy 497 CCCAGAGTATCTCAAGTGTGAGGAGAACTGTGAGTTTAACTTACCGATATGCGCAGA 556
Db 1682 CCCAGAGTATCTCAAGTGTGAGGAGAACTGTGAGTTTAACTTACCGATATGCGCAGA 1741
Qy 557 GATCCCGGAAATCACTCTCAAGAACCTTTGGAACATCACCAACCAATACACACTAC 616
Db 1742 GATCCCGGAAATCACTCTCAAGAACCTTTGGAACATCACCAACCAATACACACTAC 1801
Qy 617 ACTCAAGATCACATGATCTTATTTGGGCTTTGGAAGAACAGTCTG 663
Db 1802 ACTCAAGATCACATGATCTTATTTGGGCTTTGGAAGAACAGGATG 1848

RESULT 9

AA221228
ID AA221228 standard; cDNA; 2429 BP.
AC AA221228;

DT 22-NOV-1999 (first entry)

DE Human CG1CE long form cDNA sequence.

KW CG1CE; Best's macular dystrophy; mutation; diagnosis; detection; BMD;
age-related macular dystrophy; ss.

OS Homo sapiens.

FT Key Location/Qualifiers
FT CDS 105..1412
/*tag= a

/product= "CG1CE short form protein"

W09943695-A1.

PD 02-SEP-1999.

PF 22-FEB-1999; 99WO-US003790.

PR 25-FEB-1998; 98US-0075941P.

PR 18-DEC-1998; 98US-0112926P.

PA (MERI) MERCK & CO INC.

PA (OTUP-) UNIV UPPSALA.

PI Petrubhin K, Caskey CT, Metzker M, Wadelius C;

DR WPI; 1999-540560/45.

DR P-PSDB; AAY29954.

XX Human and mouse polynucleotides encoding CG1CE polypeptides.

XX Claim 2; Fig 4; 67bp; English.

CC The present sequence represents the human CG1CE cDNA sequence, which when
mutated is responsible for Best's macular dystrophy (BMD).

CC Polynucleotides encoding CG1CE are useful for diagnosing whether a
patient carries a mutation in the CG1CE gene. Normal and mutated CG1CE
proteins are useful for identifying activators and/or inhibitors of these

CC proteins, in order to treat BMD. The CG1CE gene offers a simpler and
cheaper method of diagnosing BMD without the need for the presence of the
CC patient. The gene may also be useful to discovering the genetic cause of
the age-related macular dystrophy

XX Sequence 2429 BP; 614 A; 694 C; 613 G; 508 T; 0 U; 0 Other;

Query Match 29.0%; Score 642.2; DB 2; Length 2429;
Best Local Similarity 99.5%; Pred. No. 7.1e-180;
Matches 644; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 17 CAGCCTGAACAAAGAGAGATGAGTTCAGCCCAATCAAGAGAGAGAGAGATGCTCA 76
Db 1405 CAGCCTGAACAAAGAGAGATGAGTTCAGCCCAATCAAGAGAGAGAGATGCTCA 1464
Qy 77 CGCTGGATATATTGGCCGCTTCTTAGGCTTCAAGTCCATATCAATCTCCAGGAGC 136
Db 1465 CGCTGGATATATTGGCCGCTTCTTAGGCTTCAAGTCCATATCAATCTCCAGGAGC 1524
Qy 137 AAATCAAGAGACCAATCACTGTCGAGGAGGAGAAATCCCTTCCAGAGGCTTCC 196
Db 1525 AAATCAAGAGACCAATCACTGTCGAGGAGGAGAAATCCCTTCCAGAGGCTTCC 1584
Qy 197 CAAAAACCAAG 256
Db 1585 CAAAAACCAAG 1644
Qy 257 GCTTAAGGCTGTGAGAGCGCTTCAAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAG 316
Db 1645 GCTTAAGGCTGTGAGAGCGCTTCAAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAG 1704
Qy 317 TGGCCACAGAGAGCGCCCTTCAAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCGC 376
Db 1705 TGGCCACAGAGAGCGCCCTTCAAGCCCACTCCATGTTCTTCCCTTAGAACCATCAGGCGC 1764
Qy 377 GTCAAGGCTTCAAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 436
Db 1765 GTCAAGGCTTCAAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1824
Qy 437 TTTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGGATGAGGCTTATGAGCA 496
Db 1825 TTTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGGATGAGGCTTATGAGCA 1884
Qy 497 CCCAGAGTATCTCAAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556
Db 1885 CCCAGAGTATCTCAAGTGTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1944
Qy 557 GATCCCGGAAATCACTCTCAAGAACCTTTGGAACATCACCAACCAATACACACTAC 616
Db 1945 GATCCCGGAAATCACTCTCAAGAACCTTTGGAACATCACCAACCAATACACACTAC 2004
Qy 617 ACTCAAGATCACATGATCTTATTTGGGCTTTGGAAGAACAGTCTG 663
Db 2005 ACTCAAGATCACATGATCTTATTTGGGCTTTGGAAGAACAGGATG 2051

RESULT 10

ADA44960
ID ADA44960 standard; cDNA; 1238 BP.

AC ADA44960;

DT 20-NOV-2003 (first entry)

DE Human polynucleotide #2.

XX Human; gene; ss; genetic disorder; genetic fingerprinting;
XX autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
XX insulin dependent diabetes mellitus; graft-versus-host disease; anaemia;
XX periodontal disease; bone fracture; cartilage damage;
XX central nervous system disorder; Alzheimer's disease;
XX Parkinson's disease; cancer; nutrition; carbon source; nitrogen source;
XX carbohydrate source.

XX OS Homo sapiens .
XX PN US2003044935-A1.
XX PD 06-MAR-2003.
XX PF 21-DEC-2000; 2000US-00746783.
XX PR 11-JUN-1997; 97US-0086236P.
PR 12-JUN-1997; 97US-0086234P.
PR 08-JUL-1997; 97US-0092115P.
PR 08-SEP-1997; 97US-0093045P.
PR 02-OCT-1997; 97US-0090100P.
PR 27-OCT-1997; 97US-00958304.
PR 07-NOV-1997; 97US-0090111P.
PR 05-JUN-1998; 98US-00092722.
PR 11-JUN-1998; 98US-00096287.
PR 17-JUN-1998; 98US-00098588.
PR 04-AUG-1998; 98US-00130189.
PR 08-SEP-1998; 98US-00149633.
PR 01-OCT-1998; 98US-00165860.
PR 04-NOV-1998; 98US-00185936.
XX PA (JACO/) JACOBS K.
PA (MCCO/) MCCOY J M.
PA (LVAL/) LA VALLIE E R.
PA (COLL/) COLLINS-RACIE L A.
PA (EVAN/) EVANS C.
PA (MERB/) MERBERG D.
PA (TREAC/) TREACY M.
PA (SPAUL/) SPAULDING V.
XX PI Jacobs K, Mccoy JM, La Vallie ER, Collins-Racie LA, Evans C;
PI Merberg D, Treacy M, Spaulding V;
XX DR MPI; 2003-521754/49.
XX DR P-SDB; ADA44961.
XX PT New polypeptides and polynucleotides having biological activities, useful
PT as nutritional sources or supplements, or for treating e.g. autoimmune
PT diseases, cancers, bone fractures or damages, or central nervous system
PT disorders.
XX PS Claim 12; Page 109; 288bp; English.
XX CC The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides can be used to express recombinant proteins
CC for analysis, characterisation or therapeutic use, as markers for tissues
CC in which the corresponding protein is expressed, as molecular weight
CC markers on Southern gels, as chromosome markers or tags to identify
CC chromosomes or to map related gene positions, to compare with endogenous
CC DNA sequences in patients to identify potential genetic disorders, as
CC probes to hybridise and discover novel related DNA sequences, as a source
CC of information to derive PCR primers for genetic fingerprinting, to raise
CC anti-protein antibodies and in gene therapy. The proteins can be used to
CC raise antibodies or to elicit another immune response, as reagents in
CC assays designed to quantitatively determine levels of the protein in
CC biological fluids, as markers for tissues in which the corresponding
CC protein is preferentially expressed and to treat autoimmune disorders
CC (e.g. multiple sclerosis, systemic lupus erythematosus, insulin dependent
CC diabetes mellitus or graft-versus-host disease), anaemias, periodontal
CC diseases, bone fractures, cartilage damage, central nervous system
CC disorders (e.g. Alzheimer's disease or Parkinson's disease) and cancers.
CC The proteins and polynucleotides are also useful as nutritional sources
CC or supplements, e.g. as carbon, nitrogen or carbohydrate sources. This
CC sequence represents a human polynucleotide of the invention.
XX SQ Sequence 1238 BP; 387 A; 336 C; 266 G; 249 T; 0 U; 0 Other;

Query Match 23.1%; Score 511.8; DB 8; Length 1238;
Best Local Similarity 98.0%; Pred. No. 3.2e-141;
Matches 634; Conservative 0; Mismatches 2; Indels 11; Gaps 11;

QY 17 CAGCCTGAACAAAGAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 76
|||
Db 223 CAGCCTGAACAAAGAGAGATGAGATTCCAGCCCAATCAGAGAGACGAGAGATGCTCA 281
|||
QY 77 CGCTGGCATCATTTGGCCGCTTCTCTAGGCTGACAGTCCCATGATCCATCTCCCAAGGC 136
|||
Db 282 CGCTGGCATCATTTGGG-CGCTTCTTAGGCTGACAGTCCCATGATCCATCTCCCAAGGC 340
|||
QY 137 AAACCTCAAGGACCAAACTACTGTGGCCCAAGAGGAATCCCTTCTCAGAGAGGCTTGGC 196
|||
Db 341 AAACCTCAAGGAGC-AACTACTGTGGCCCAAGAGGAATCCCTTCTCAGAGAGGCTTGGC 399
|||
QY 197 CAAAAACCAAGGACGACCAACAGAACGTTAAGGGCCAGAGAACCAAGGCTTGA 256
400 CAAAAACCAAGGACGACCAACAGAACGTTAAGGGCCAGAGAACCAAGGCTTGA 458
|||
QY 257 GCTTAAAGCTGTGAGAGCCTTCAAGTCTGCCCCACTGTATCAGAGGCCAGGCTTACAG 316
459 GCTTAAAGCTGTGAGAGCCTTCAAGTCTGCCCCACTGTATCAGAGGCCAGGCTTACAG 517
|||
QY 317 TGCCCCACAGAGCCCTTCAAGCTCCCACTCCCATGTTCTTCCCTTGAACCATCAGGCC 376
518 TGCCCCACAGAGC-CCCTCAGCCCACTCCCATGTTCTTCCCTTGAACCATCAGGCC 576
|||
QY 377 GTCMAAGCTTCAAGTGTCAAGGATAGACCAACCAAGGCTTAAAGCTTGAAG 436
577 GTCMAAGCTTCAAGGATAGACCAACCAAGGCTTAAAGCTTGAAG 635
|||
QY 437 TTCTGGGGCCCAAGAAAGTTTGAATTTGCTCTCAGAGAGGAGTGGGCTTGAATGAGCA 496
636 TTCTGGGGCCCAAG-AAAGTTTGAATTTGCTCTCAGAGAGGAGTGGGCTTGAATGAGCA 694
|||
QY 497 CCCAGAAATCTCAAGTGAAGGAGAAACTGTGAGTTTAACTGACGATATGCCAGA 556
695 CCCAGAAATCTCAAGTGAAGGAGAAACTGTGAGTTTAACTGACGATATGCCAGA 753
|||
QY 557 GATCCCCGAAATCACCTCAAGACCTTTGGAACATCCAAACCAATCAGACTAC 616
754 GATCCCCGAAATCA-CTCAAGACCTTTGGAACATCCAAACCAATCAGACTAC 812
|||
QY 617 ACTCAAGATCAGATGATCTTATTTGGGCTTGAACAGGCTG 663
813 ACTCAAGATCAGATGATCTTATTTGGGCTTGAACAGGAGT 858
|||
Db
RESULT 11
ID ABA14556 standard; DNA; 7108 BP.
XX ABA14556;
XX AC
XX 23-JAN-2002 (first entry)
XX DT
XX XX
XX XX
XX Human nervous system related polynucleotide SEQ ID NO 6887.
XX DE
XX XX
KW Human; noctropic; neuroprotective; cyrostatic; dermatological; virocidic;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antisticking; antianaemic; antirheumatic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antineuroinflammatory;
KW antiallergic; antidiabetic; antituber; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX OS Homo sapiens.
XX OS
XX PN WO200159063-A2.
XX PD 16-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001334.
XX PR 31-JAN-2000; 2000US-0179065P.

PS Disclosure; SEQ ID NO 6887; 1701bp + Sequence Listing; English.

XX
CC The invention relates to novel genes (ABAI1004-ABAI21534) and proteins
CC (ABAI678-ABAI8001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC / (e) neurological diseases such as cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at http://pub.wipo.int/pub/published_pcr_sequences

XX
SQ Sequence 7108 BP; 1632 A; 1882 C; 1989 G; 1605 T; 0 U; 0 Other;

Query Match 22.3%; Score 494; DB 5; Length 7108;
Best Local Similarity 99.6%; Pred. No. 1.8e-135;
Matches 516; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1 ATCTCCGTTCTTTCCAGCGCTGACCAAGAGAGATGAGTTCACGCCCAATCAGAGG 60
DB 6593 ATCTCCGTTCTTTCCAGCGCTGACCAAGAGAGATGAGTTCACGCCCAATCAGAGG 6552

QY 61 ACAGAGAGATGTCACGCTGCGCATTCATTTGGCCGCTTCTTAGGCTCAGTCCCATATC 120
DB 6653 ACAGAGAGATG-TCAAGCTGCGCATTCATTTGGCCGCTTCTTAGGCTCAGTCCCATATC 6711

QY 121 ACCATCTCCAGGCGCAAACTCAAGACCAAACTACTGTGTGCGCCCAAGAGGAAATCCCTTC 180
DB 6712 ACCATCTCCAGGCGCAAACTCAAGACCAAACTACTGTGTGCGCCCAAGAGGAAATCCCTTC 6771

QY 181 TCCACGAGGGGCTGCGCCCAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 6772 TCCACGAG-GGCTGCGCCCAAAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 6830

QY 241 ACAAGAGAGGCTGAGAGCTTAAGGCTGTGAGCGCTTCAAGTCTGCGCCCACTGATCAGA 300
DB 6831 ACAAGAGAGGCTGAGAGCTTAAGGCTGTGAGCGCTTCAAGTCTGCGCCCACTGATCAGA 6890

QY 301 GGGCAGGCTACTACAGTGCAGCCCAAGAGAGGCGCCCTCAGCGCCCATCTCCATGTTCTCCCC 360
DB 6891 GGGCAGGCTACTACAGTGCAGCCCAAGAGAGGCGCCCTCAGCGCCCATCTCCATGTTCTCCCC 6950

QY 361 TAAAGCATATGAGGCGCTCAAGGTTCAAGTGTCAAGGCTATGAGACCAAGAGCAAAA 420
DB 6951 TAAAGCATATGAGGCGCTCAAGGTTCAAGTGTCAAGGCTATGAGACCAAGAGCAAAA 7010

QY 421 GCTTAAAGACTGTGAGTTCGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATG 480
DB 7011 GCTTAAAGACTGTGAGTTCGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCATG 7070

QY 481 GGGCCTTGATGAGACCAAGATATCTCAAGTGAAG 518
DB 7071 GGGCCTTGATGAGACCAAGATATCTCAAGTGAAG 7108

RESULT 12
ID ACH39717 standard; cDNA; 459 BP.
XX ACH39717;
AC ACH39717;
XX
DT 13-OCT-2003 (first entry)
XX Human foetal brain cDNA #1084.
XX

KM Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KM genome mapping; biodiversity; genetic disorder.

XX
OS Homo sapiens.
XX
PN US2003073623-A1.
XX
PD 17-Apr-2003.
XX
PF 30-JUL-2001; 2001US-00918995.
XX
PR 30-JUL-2001; 2001US-00918995.
XX
PA (DRMA/) DRMANAC R T.
PA (LABA/) LABAT I.
PA (STAC/) STACHE-CRAIN B.
PA (DICK/) DICKSON M C.
PA (JONE/) JONES L W.
XX
PI Dermanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;
XX
DR WPI; 2003-615964/58.
XX
XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.

PS
XX Claim 1; SEQ ID NO 26929; 44bp; English.

XX
CC The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostic as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX
SQ Sequence 459 BP; 134 A; 90 C; 122 G; 111 T; 0 U; 2 Other;

Query Match 18.2%; Score 404.2; DB 8; Length 459;
Best Local Similarity 96.4%; Pred. No. 2e-107;
Matches 423; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 1333 CGTGGAGGAAAGTAACTCCCTTTCTGATTTCAAGCAGTACTTTACAGGGGTCAG 1392
DB 22 CGAGGAATTCGTGTAATCTCCCTTTCTGATTTCAAGCAGTACTTTAC-GGTCAG 80

QY 1393 AACACGACGATATTAATGATTAAGAACTTAAGAGGCAACATTTCAATCTTGGCTTAGG 1452
DB 81 AACACGACGATATTAATGATTAAGAACTTAAGAGGCAACATTTCAATCTTGGCTTAGG 140

QY 1453 CTAAAGACAGGAATTTGGCAAACTCTGTGGCTTGTCAAGCAAGATGTTCAATTTTAAG 1512
DB 141 GCTAAGACAGGAATTTGGCAAACTCTGTGGCTTGTCAAGCAAGATGTTTAATTTAAG 200

QY 1513 AATCTGTCTTGGGCTGGGTGTGAGGCAAGTAAATCAAGAGGTCAAGAGTTTGAAG 1572
DB 201 AATCTGTCTTGGGCTGGGTGTGAGGCAAGTAAATCAAGAGGTCAAGAGTTTGAAG 260

QY 1573 CAACCTGGCCAAATGATGAAACCCCATCTTACCAAAAAATTAACAATCACTGGCG 1632
DB 261 CAACCTGGCCAAATGATGAAACCCCATCTTACCAAAAAATTAACAATCACTGGCG 320

CC or their complements that are differentially expressed in cancer and
CC other proliferative disorders. The combination is useful in detecting
CC changes in expression of genes encoding proteins that are associated with
CC senescence and in diagnosing, staging, treating, or monitoring the
CC progression or treatment of subjects with proliferative diseases such as
CC cancer (e.g. adenocarcinoma, leukemia, lymphoma, melanoma, myeloma,
CC sarcoma, leiomyosarcoma, cancer of the adrenal gland, bladder, bone,
CC bone marrow, brain, breast, cervix, colon, esophagus, gall bladder,
CC ganglia, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid
CC gland, penis, prostate, salivary glands, skin, small intestine, spleen,
CC stomach, testis, thymus, thyroid and uterus. The present sequence
CC represents cDNA of genes that are upregulated in senescent cells
XX

Sequence 1326 BP; 338 A; 328 C; 315 G; 345 T; 0 U; 0 Other;

Query Match 15.3%; Score 339.4; DB 7; Length 1326;
Best Local Similarity 99.7%; Pred. No. 7.2e-90;
Matches 340; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1876 AGCAGATTCCTAACCTGCTTCTTATATGGGAGTGGTCCAGCCAGGTCCTCACTGTGT 1935
DB 1326 AGCAGATTCCTAACCTGCTTCTTATATGGGAGTGGTCCAGGTCCTCACTGTGT 1267
QY 1936 GTAACACGAGAGGAGCACTGATCCAGTCACAGCCATACAGCTGTCCACATGAAAGCTG 1995
DB 1266 GTAACACGAGAGGAGCACTGATCCAGTCACAGCCATACAGCTGTCCACATGAAAGCTG 1207
QY 1996 TCTCAACAAGCCTGATCAATATGCTTAAATAGATTAATAATCCCAAGCTACTTCAAG 2055
DB 1206 TCTCAACAAGCCTGATCAATATGCTTAAATAGATTAATAATCCCAAGCTACTTCAAG 1147
QY 2056 CCTTATGCTTATTTATTTATTAATAAAGTGAAGCTGAGCACTGGAACATTT 2115
DB 1146 CCTTATGCTTATTTATTTATTTATTAATAAAGTGAAGCTGAGCACTGGAACATTT 1087
QY 2116 AACTCAGACTCTGATTCAGAGTCGGGAAACCTTATGTTCTATGTAATCCCAAGCAAGCA 2175
DB 1086 AACTCAGACTCTGATTCAGAGTCGGGAAACCTTATGTTCTATGTAATCCCAAGCAAGCA 1027
QY 2176 CACCTTAGTATCTGCCCCAAACTATATAGTTTAATAATAC 2216
DB 1026 CACCTTAGTATCTGCCCCAAACTATATAGTTTAATAATAC 986

RESULT 15

AAC20667 standard; cDNA; 356 BP.

XX AAC20667;
AC AAC20667;
XX 06-OCT-2000 (first entry)
XX Human secreted protein 5' EST, SEQ ID NO: 24742.
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping; ss.
XX Homo sapiens.
OS
XX EP1033401-A2.
PN
XX 06-SEP-2000.
PD
XX 21-FEB-2000; 2000EP-00200610.
PF
XX 26-FEB-1999; 99US-0122487P.
PR
XX (GENEST) GENSET.
PA
XX Dumas Mline Edwards J, Duclert A, Giordano J;
PI
XX WPI; 2000-500381/45.
XX

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 24742; 71bp + Sequence Listing; English.
XX

PS The present sequence is one of a large number of 5' ESTs derived from
PS mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNA or poly(A) RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX

Sequence 356 BP; 103 A; 69 C; 90 G; 88 T; 0 U; 6 Other;

Query Match 14.4%; Score 318; DB 3; Length 356;
Best Local Similarity 95.8%; Pred. No. 7.8e-84;
Matches 342; Conservative 4; Mismatches 8; Indels 3; Gaps 2;

QY 1331 CACGTGGAGAGAGTGAATCTTCCCTTTCTGATTTCTCAAGCACTTCTTACGGGTC 1390
DB 1 CACGTGGAGAGAGTGAATCTTCCCTTTCTGATTTCTCAAGCACTTCTTAC--GGTC 59
QY 1391 AGAACACGAGCACTTATATGATTAAGCTTAAGGCAACATTCANTCTGCTCTA 1450
DB 60 AGAACACGAGCACTTATATGATTAAGCTTAAGGCAACATTCAGTCTTGCTCTA 119
QY 1451 GGCTAAGACAGAACTTGGCAACATCTGTGGCTGTCAGCAAGAGTGTCAATTTA 1510
DB 120 GGGYMAAGACAGAACTTGGCAACATCTGTGGCTGTCAGCAAGAGTGTCAATTTA 179
QY 1511 AGAATCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATCAAGAGAGTCAGAGTTTGA 1570
DB 180 AGAATCTTGTCTTGGGCTGGGTGTGAGGCAAGTGAATCAAGAGAGTCAGAGTTTGA 239
QY 1571 ACCAAGCTGGCAACATGATGAAGCAACCCATCTTACCAAAAAAATAC--AATCAGCTG 1628
DB 240 ACCAAGCTGGCAACATGATGAAGCAACCCATCTTACCAAAAAAATACAAATCAGCTG 299
QY 1629 GCCGTGCTGGTGTGCTGTAGTCCCAAGCAGAGAGTTGAGGAGAAATTTGTTGAA 1685
DB 300 GCCGTGCTGGTGTGCTGTAGTCCCAAGCAGAGAGTTGAGGAGAAATTTGTTGAA 356

Search completed: March 25, 2004, 03:23:35
Job time : 861 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 25, 2004, 03:09:19 ; Search time 162 Seconds

(without alignments)
7591.180 Million cell updates/sec

Title: US-09-622-964-1_COPY_13900_16115

Perfect score: 2216
Sequence: 1 atctcctgtctcttcacgc.....ctaagttaataaatac 2216

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 27475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents NA: *
1: /cgn2_6/ptodaca/2/ina/5A COMB.seq: *
2: /cgn2_6/ptodaca/2/ina/5B COMB.seq: *
3: /cgn2_6/ptodaca/2/ina/6A COMB.seq: *
4: /cgn2_6/ptodaca/2/ina/6B COMB.seq: *
5: /cgn2_6/ptodaca/2/ina/PTUS COMB.seq: *
6: /cgn2_6/ptodaca/2/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	133.2	6.0	65042	US-09-784-316-3	Sequence 3, Appl1
2	123	5.6	7720	US-09-318-448-5	Sequence 5, Appl1
3	121.6	5.5	36159	US-09-749-588-3	Sequence 3, Appl1
4	119.4	5.4	49312	US-09-671-317-485	Sequence 485, App
5	117.2	5.3	2867	US-09-402-532-38	Sequence 38, Appl1
6	117.2	5.3	21234	US-09-810-671-3	Sequence 3, Appl1
7	117.2	5.3	21234	US-10-109-854-3	Sequence 42, Appl1
8	117	5.3	44848	US-09-435-739-42	Sequence 1, Appl1
9	117	5.3	162450	US-09-345-882-1	Sequence 1, Appl1
10	116.8	5.3	90541	US-09-759-359A-3	Sequence 1, Appl1
11	116.8	5.3	92139	US-09-918-686-1	Sequence 1, Appl1
12	116.2	5.2	62804	US-09-800-960-3	Sequence 3, Appl1
13	116.2	5.2	62804	US-10-096-960-3	Sequence 3, Appl1
14	115.8	5.2	18853	US-09-820-005-3	Sequence 3, Appl1
15	115.4	5.2	2296	US-09-851-896-11	Sequence 11, Appl1
16	115.2	5.2	99500	US-09-798-096-10	Sequence 10, Appl1
17	115.2	5.2	188575	US-09-426-290-1	Sequence 1, Appl1
18	114	5.1	246240	US-08-724-394A-20	Sequence 20, Appl1
19	114	5.1	246240	US-08-724-394A-21	Sequence 21, Appl1
20	114	5.1	246240	US-08-724-394A-22	Sequence 22, Appl1
21	113.8	5.1	2236	US-08-829-525-23	Sequence 23, Appl1
22	113.8	5.1	2236	US-08-609-583A-23	Sequence 23, Appl1
23	113.8	5.1	2236	US-08-937-399-23	Sequence 23, Appl1
24	113.8	5.1	2236	US-09-310-367-23	Sequence 23, Appl1
25	113.8	5.1	2236	US-09-032-337-23	Sequence 23, Appl1
26	113.8	5.1	2236	US-09-464-231-23	Sequence 23, Appl1
27	113.8	5.1	53332	US-09-801-861-3	Sequence 3, Appl1

C	28	113.6	5.1	72604	4	US-09-268-992-7	Sequence 7, Appl1
C	29	113.6	5.1	72604	4	US-09-657-474-7	Sequence 7, Appl1
C	30	113.6	5.1	87350	4	US-08-781-891-79	Sequence 79, Appl1
C	31	113.6	5.1	87350	4	US-09-618-166-79	Sequence 79, Appl1
C	32	113.6	5.1	87353	4	US-09-791-211-3	Sequence 3, Appl1
C	33	113	5.1	18596	3	US-09-318-448-11	Sequence 11, Appl1
C	34	113	5.1	21234	4	US-09-810-671-3	Sequence 3, Appl1
C	35	113	5.1	21234	4	US-10-109-854-3	Sequence 3, Appl1
C	36	112.8	5.1	21721	4	US-09-269-939A-41	Sequence 41, Appl1
C	37	112.8	5.1	22976	4	US-09-269-939A-19	Sequence 19, Appl1
C	38	112.8	5.1	23187	4	US-09-499-522-1	Sequence 1, Appl1
C	39	112.8	5.1	64467	4	US-09-803-671B-3	Sequence 3, Appl1
C	40	112.8	5.1	70000	4	US-09-851-896-3	Sequence 3, Appl1
C	41	112.8	5.1	162450	4	US-09-345-882-1	Sequence 1419, Ap
C	42	112.6	5.1	2932	4	US-09-016-434-1419	Sequence 5, Appl1
C	43	112	5.1	6235	3	US-09-305-384-5	Sequence 5, Appl1
C	44	112	5.1	6235	4	US-09-525-160B-6	Sequence 6, Appl1
C	45	112	5.1	6679	3	US-09-305-384-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-09-784-316-3/C
; Sequence 3, Application US/09784316
; Patent No. 6461843
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; FILE REFERENCE: CL001139
; CURRENT APPLICATION NUMBER: US/09/784,316
; CURRENT FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 65042
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(65042)
; OTHER INFORMATION: n = A,T,C or G
US-09-784-316-3

Query Match 6.0%; Score 133.2; DB 4; Length 65042;
Best Local Similarity 70.6%; Pred. No. 5,1e-31;
Matches 209; Conservative 0; Mismatches 78; Indels 9; Gaps 2;

QY	1527	CTGGTGTGAGGCAAGTGAATCAGAGGTCAGAGTTTGAACCACTGGCCNAACA	1586
DB	35972	CGGAGAGCGAGGCGGGCGGATCACTGAGGTCCAGAGTCCAGACCAAGCTGGCCNAACA	35913
QY	1587	TGATGAAGCCCTCTCTCAAAAAATTAACAATCAGTGGCCGTGGTGGCTG	1646
DB	35912	TGATGAAGCCCTCTCTCTCAAAAAATTAACAATCAGTGGCCGTGGTGGCTG	35853
QY	1647	TGATCCCAAC---GCAGAGGTTGAGGAGAAATGCTTGAACCCAGAGAGTGTGCTT	1702
DB	35852	TATCCACACTACTCAAGAGGCTGGGACAGAAATCACTTGAACCCAGAGAGTGTGCTT	35793
QY	1703	GCAG---TGAGATTGAGCACTGCAATTCACCTGGGCGACGAGAGTGTGCTCA	1757
DB	35792	GCAGTAGCTGAGATTGTCCACTGCACTCCAGCTGGGCGAAGAGAGTGTGCTCT	35733
QY	1758	AAAAAAAAAAAAAAAAAGATGCTGCAACTTTGCTCTCACTGCAACTTTGCT	1813
DB	35732	CAAAAAAAAAAAAAAAAAAGTATGCTGTAAGTAAACATTCAGT	35677

RESULT 2

```

US-09-318-448-5/c
; Sequence 5, Application US/09318448
; Patent No. 6210950
; GENERAL INFORMATION:
; APPLICANT: Johnson, William G.
; APPLICANT: Stemroos, Edward S.
; TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING
; FILE REFERENCE: 601-1-057
; CURRENT APPLICATION NUMBER: US/09/318,448
; CURRENT FILING DATE: 1999-05-25
; NUMBER OF SEQ. ID NOS.: 46
; SOFTWARE: PatentIn Ver. 2.0
; SEQ. ID NO. 5
; LENGTH: 7720
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-318-448-5

```

Query Match	5.6%	Score 123;	DB 3;	Length 7720;
Best Local Similarity	71.0%	Pred. No. 2..1e-28;		
Matches 208;	Conservative	0;	Mismatches 75;	Indels 1

QY	1536	GAGGCAAGTGAATCA	CAGAGGCGAGACTTTG	GAAACCA	CTGGCCAA	CATGATGAAC	1595
Db	3007	GAGTGGGTGATCACT	GAGGTCA	GAATTTGA	GCACCTGGCCAA	CATAGGAAC	2948
QY	1596	CCCATCTTACCAAAAA	AAATCAAA	TACAGCTGCGCC	TGTGTGTG	CTGTAGTCCCA	1654
Db	2947	CTGTCTCTACAAAA	AAATTCAAAA	TAGAGCTGGGTG	CTGTCTGTGCTGTAGTCC	CAAG	2888
QY	1655	--ACCGACGAGGTT	GAG--GGAGA	ATTTGCTTGA	CCGAGAGGTG	GTGTGACAGT--	1708
Db	2887	CTACCTGGAGGCT	GTAGGCGAGAG	AATCACTTGA	TCCAGAGGTG	AGGTTACAGT	2828
QY	1709	---AGATTGACCA	CTGCAATCGA	CCCTGGCGCA	CGAGAGTGA	CACTGTCTCAAAAA	1765
Db	2827	CGAAGATTGTCC	CACTTGTGCA	CTGTGCTG	GCACAGGCG	CACTCACTTCAAAAA	2768
QY	1766	AAAAAAGATCG	TCTCAACCTT	GCCCTCTA	CTGCACATTT	TGGTATTG	1818
Db	2767	AAAAAAGTTAT	ATAAAATGTTT	TAGTGGGGGG	GTGCA	GTGTGTTTG	2715

```

RESULT 3
US-09-749-588-3/C
; Sequence 3, Application US/09749588
; Patent No. 6423521
; GENERAL INFORMATION:
; APPLICANT: CHANDRAMOULISWARAN, Ishwar et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROPF
; FILE REFERENCE: C1001068
; CURRENT APPLICATION NUMBER: US/09/749,588
; CURRENT FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 36159
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(36159)
; OTHER INFORMATION: n = A,T,C or G
US-09-749-588-3

```

Query Match	5.5%	Score 121.6;	DB 4;	Length 36159;
Best Local Similarity	78.1%;	Pred. No. 1.9e-27;		
Matches 200; Conservative	0;	Mismatches 44;		

1529 GGGTGTGAGGCAAGTGAATCACAGGAGTCAGAGGTTTGAGACCAACTGGCCACATG 1588

Accession	Sequence	Position
Db	11731 GGAGGCGAGAGCGCTGGATCATCTGAACTCAGAGTTTGAAGCAGAGCCTGGCCAATG	11672
Qy	1589 ATGAAGCCCACTCTCTATCAAAAAAAAAATACAAATCAGCTGGCCGTCGTGTGTG--CTTG	1646
Db	11671 GTGAAGCCCACTCTCTATCAAAAAAAAAATACAGCCAGGTGTGTGTGCGGGCACCTG	11612
Qy	1647 TAGTCCCA--ACGCAGAGAGTTGAGG--GGAATATTGCTGTAACCCAGAGATGTGTG	1701
Db	11611 TAAATCCAGCTACTCAGAGAGCTGAGGCGAGAGAAATTGCTTGAACCCAGAGAACGAGGT	11552
Qy	1702 TGCAGTAGG-----ATTGAGCAACTGCAATCCAGCCTGGGCGACGAGTGAAGACTGTCTC	1756
Db	11551 TGCAGTAGGCTGACACAGTGCCTCCTCATTCAAGCTGGGTACAGAGTTAAGACTCTGTCTC	11492
Qy	1757 AAAAAAAAAAAAAAAAAA	1772
Db	11491 TCAAAAAAAAAAAAAA	11476

RESULT 4
US-09-671-317-485

```

? Sequence 488, Application US/09671317
? Patent No. 6528260
? GENERAL INFORMATION:
? APPLICANT: Blumenfeld, Marta
? APPLICANT: Chumakov, Ilya
? APPLICANT: Bouguerelet, Lydie
? APPLICANT: Cohen, Amick
? TITLE OF INVENTION: BIALLELIC MARKERS RELATED TO GENES INVOLVED IN DRUG METABOLISM
? FILE REFERENCE: 62, US3, CIP
? CURRENT APPLICATION NUMBER: US/09/671,317
? CURRENT FILING DATE: 2000-09-27
? PRIOR APPLICATION NUMBER: US 09/536,118
? PRIOR FILING DATE: 2000-03-23
? PRIOR APPLICATION NUMBER: PCT/IB00/00403
? PRIOR FILING DATE: 2000-03-24
? PRIOR APPLICATION NUMBER: US 60/126,269
? PRIOR FILING DATE: 1999-03-25
? PRIOR APPLICATION NUMBER: US 60/131,361
? PRIOR FILING DATE: 1999-04-30
? NUMBER OF SEQ ID NOS: 977
? SOFTWARE: Patent.pm
? SEQ ID NO 485

```

```

1 TYPE: DNA
2 ORGANISM: Homo sapiens
3 FEATURE:
4 NAME/KEY: misc_feature
5 LOCATION: 5466..7466
6 OTHER INFORMATION: 5'regulatory region
7 NAME/KEY: exon
8 LOCATION: 7467..7725
9 OTHER INFORMATION:
10 NAME/KEY: exon 1
11 LOCATION: 20256..20355
12 OTHER INFORMATION: exon 2
13 NAME/KEY: exon
14 LOCATION: 36905..36975
15 OTHER INFORMATION: exon 3
16 NAME/KEY: exon
17 LOCATION: 45167..45248
18 OTHER INFORMATION: exon 4
19 NAME/KEY: exon
20 LOCATION: 45728..45965
21 OTHER INFORMATION: exon 5
22 NAME/KEY: misc_feature
23 LOCATION: 45966..49312
24 OTHER INFORMATION: 3'regulatory region
25 NAME/KEY: allele
26 LOCATION: 7564
27 OTHER INFORMATION: 10-286-289 : polymorphic base G or C
28 NAME/KEY: allele

```

```
LOCATION: 7619
OTHER INFORMATION: 10-286-345 : polymorphic base A or T
NAME/KEY: allele
LOCATION: 7649
OTHER INFORMATION: 10-286-375 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 17258
OTHER INFORMATION: 12-425-57 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 21590
OTHER INFORMATION: 12-421-135 : insertion of T
NAME/KEY: allele
LOCATION: 21595
OTHER INFORMATION: 12-421-140 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 36971
OTHER INFORMATION: 10-523-232 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 45214
OTHER INFORMATION: 10-289-201 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 45741
OTHER INFORMATION: 10-290-37 : polymorphic base C or T
NAME/KEY: allele
LOCATION: 46029
OTHER INFORMATION: 10-290-326 : polymorphic base A or G
NAME/KEY: allele
LOCATION: 46032
OTHER INFORMATION: 10-290-328 : deletion of G
NAME/KEY: primer bind
LOCATION: 7276..7294
OTHER INFORMATION: 10-286.pu
NAME/KEY: primer bind
LOCATION: 7676..7694
OTHER INFORMATION: 10-286.rp complement
NAME/KEY: primer bind
LOCATION: 16839..16856
OTHER INFORMATION: 12-425.rp
NAME/KEY: primer bind
LOCATION: 17297..17314
OTHER INFORMATION: 12-425.pu complement
NAME/KEY: primer bind
LOCATION: 21456..21474
OTHER INFORMATION: 12-421.pu
NAME/KEY: primer bind
LOCATION: 21886..21906
OTHER INFORMATION: 12-421.rp complement
NAME/KEY: primer bind
LOCATION: 36740..36758
OTHER INFORMATION: 10-523.pu
NAME/KEY: primer bind
LOCATION: 36997..37015
OTHER INFORMATION: 10-523.rp complement
NAME/KEY: primer bind
LOCATION: 45020..45037
OTHER INFORMATION: 10-289.pu
NAME/KEY: primer bind
LOCATION: 45413..45432
OTHER INFORMATION: 10-289.rp complement
NAME/KEY: primer bind
LOCATION: 45705..45724
OTHER INFORMATION: 10-290.pu
NAME/KEY: primer bind
LOCATION: 46104..46123
OTHER INFORMATION: 10-290.rp complement
NAME/KEY: primer bind
LOCATION: 7545..7563
OTHER INFORMATION: 10-286-289.mis
NAME/KEY: primer bind
LOCATION: 7565..7583
OTHER INFORMATION: 10-286-289.mis complement
NAME/KEY: primer bind
LOCATION: 7600..7618
OTHER INFORMATION: 10-286-345.mis
NAME/KEY: primer bind
LOCATION: 7620..7638
OTHER INFORMATION: 10-286-345.mis complement
NAME/KEY: primer bind
LOCATION: 7630..7648
OTHER INFORMATION: 10-286-375.mis
NAME/KEY: primer bind
LOCATION: 7650..7668
OTHER INFORMATION: 10-286-375.mis complement
NAME/KEY: primer bind
LOCATION: 17239..17257
OTHER INFORMATION: 12-425-57.mis
NAME/KEY: primer bind
LOCATION: 17259..17277
OTHER INFORMATION: 12-425-57.mis complement
NAME/KEY: primer bind
LOCATION: 21576..21594
OTHER INFORMATION: 12-421-140.mis
NAME/KEY: primer bind
LOCATION: 21596..21614
OTHER INFORMATION: 12-421-140.mis complement
NAME/KEY: primer bind
LOCATION: 36952..36970
OTHER INFORMATION: 10-523-232.mis
NAME/KEY: primer bind
LOCATION: 36972..36990
OTHER INFORMATION: 10-523-232.mis complement
NAME/KEY: primer bind
LOCATION: 45195..45213
OTHER INFORMATION: 10-289-201.mis
NAME/KEY: primer bind
LOCATION: 45215..45233
OTHER INFORMATION: 10-289-201.mis complement
NAME/KEY: primer bind
LOCATION: 45722..45740
OTHER INFORMATION: 10-290-37.mis
NAME/KEY: primer bind
LOCATION: 45742..45760
OTHER INFORMATION: 10-290-37.mis complement
NAME/KEY: primer bind
LOCATION: 46010..46028
OTHER INFORMATION: 10-290-326.mis
NAME/KEY: primer bind
LOCATION: 46030..46048
OTHER INFORMATION: 10-290-326.mis complement
NAME/KEY: primer bind
LOCATION: 7552..7576
OTHER INFORMATION: 10-286-289.probe
NAME/KEY: misc binding
LOCATION: 7607..7631
OTHER INFORMATION: 10-286-345.probe
NAME/KEY: misc binding
LOCATION: 7637..7661
OTHER INFORMATION: 10-286-375.probe
NAME/KEY: misc binding
LOCATION: 17246..17270
OTHER INFORMATION: 12-425-57.probe
NAME/KEY: misc binding
LOCATION: 21583..21607
OTHER INFORMATION: 12-421-140.probe
NAME/KEY: misc binding
LOCATION: 36959..36983
OTHER INFORMATION: 10-523-232.probe
NAME/KEY: misc binding
LOCATION: 45202..45226
OTHER INFORMATION: 10-289-201.probe
NAME/KEY: misc binding
LOCATION: 45725..45753
OTHER INFORMATION: 10-290-37.probe
NAME/KEY: misc binding
LOCATION: 46017..46041
OTHER INFORMATION: 10-290-326.probe
```

US-09-671-317-485

Query Match
 Best Local Similarity 5.4%; Score 119.4; DB 4; Length 49312;
 Matches 205; Conservative 0; Mismatches 56; Indels 12; Gaps 4;

QY 1528 TGGGTGTGAGCAATGATCAAGAGCTGAGGTTTGAACCACTTGCCCAAT 1587
 DB 16251 TGGAGGCTAAGCGGGGAGATCACTTGAGGTGAGGTTCAAGACCAGCTTGCCCAAT 16310
 QY 1588 GATGAACCCCATCTCTCAACAAAAATACAATAGCTGGCCGTGGTGG--TGTGCTT 1645
 DB 16311 GTGTAACCCCATCTCTCTCAACAAAAATACAATAGCTGGCCGTGGTGGTGGT 16370
 QY 1646 GTAATCCCA--ACGAGAGAGGTGAG--GGAATTTGCTTGAACCCAGAGGTGGTGG 1700
 DB 16371 GTAATCCCACTCTAGAGAGTTGAGGAGAGAAATGCTTGAACCCAGAGGTGAGG 16430
 QY 1701 TTGCAG----TGAGTTGAGCACTGCAATCCAGCTGGGCGAGAGGAGACTGTCT 1755
 DB 16431 TTGCAGTGAAGCTGAGATCAATGCACTGCACTCCAGCTGGGCGAGAGGAGACTGTCT 1755
 QY 1756 CAAAAAAGAGATGCTCAACTT 1788
 DB 16491 CTCTTAAAAAACAATCTTATAGCTATCTT 16523

RESULT 5

US-09-402-532-38
 ; Sequence 38, Application US/09402532
 ; Patent No. 6498019
 ; GENERAL INFORMATION:
 ; APPLICANT: Taniyama, Yoshio
 ; TITLE OF INVENTION: NOVEL PROTEIN, ITS PRODUCTION AND USE
 ; FILE REFERENCE: 2456USOP
 ; CURRENT APPLICATION NUMBER: US/09/402,532
 ; EARLIER FILING DATE: 1999-10-04
 ; EARLIER APPLICATION NUMBER: PCT/JP98/01643
 ; EARLIER FILING DATE: 1998-04-09
 ; EARLIER APPLICATION NUMBER: JP 10-010289
 ; EARLIER FILING DATE: 1998-01-22
 ; EARLIER APPLICATION NUMBER: JP 9-184885
 ; EARLIER FILING DATE: 1997-07-10
 ; EARLIER APPLICATION NUMBER: JP 9-093355
 ; NUMBER OF SEQ ID NOS: 43
 ; SOFTWARE:
 ; SEQ ID NO 38
 ; LENGTH: 2867
 ; TYPE: DNA
 ; ORGANISM: Unknown
 ; FEATURE:
 ; OTHER INFORMATION: Unsure (5' flanking promoter sequence of genomic DNA)
 ; US-09-402-532-38

Query Match
 Best Local Similarity 5.3%; Score 117.2; DB 4; Length 2867;
 Matches 195; Conservative 0; Mismatches 43; Indels 12; Gaps 4;

QY 1537 AGGCAAGTATCAACAGAGCTCAAGAGTTTGAACCACTGGCCCAAGATGAAC 1596
 DB 104 AGGTGGTGTATCACTGAGAGTCAAGAGTTTGAACCACTGGCCCAAGATGAAC 163
 QY 1597 CCAATCTTACCAAAAAATACAATCACTGGCCGTGGT--GTGTGCTTGTAGTCCA 1654
 DB 164 CCGTCTCTCTAAAAAATACAATCACTGGCCGTGGTGGTGGTGGTGGTGGTGGT 223
 QY 1655 ---ACGAGAGAGTGGAG--GGAATTTGCTTGAACCCAGAGAGTGGTGGTGGTGA 1709
 DB 224 GCAATTCAGAGAGCTGAGGAGAGAAATCGCTTGAACCCAGAGAGAGTGGTGGTGA 283
 QY 1710 ---GATTGACCACTGCAATCCAGCTGGGCGAGAGGAGAGTGTCTCAAAAAA 1764

DB 284 GCGGAGATCAGCCCACTGCTCTCCAGCTGGTGACAGCAAAACTGTGTCAAAAA 343
 QY 1765 AAAAAAAGG 1774
 DB 344 AAAAAAAGG 353

RESULT 6

US-09-810-671-3/c
 ; Sequence 3, Application US/09810671
 ; Patent No. 6455291
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: C1000758
 ; CURRENT APPLICATION NUMBER: US/09/810,671
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 21234
 ; TYPE: DNA
 ; ORGANISM: Human
 ; US-09-810-671-3

Query Match
 Best Local Similarity 5.3%; Score 117.2; DB 4; Length 21234;
 Matches 198; Conservative 0; Mismatches 48; Indels 12; Gaps 4;

QY 1529 GGGTGTGAGGCAAGTATCAAGAGCTCAGAGTTTGAACCACTGGCCCAATG 1588
 DB 11068 GGAGGCGAGGACAGAGATCACTTGAGGTGAGAGTTGAGACAGCTGGCTAATG 11009
 QY 1589 ATGAACCCCATCTCTCAACAAAAATACAATCACTGGCCGTGGTGGTGGT--CCTG 1646
 DB 11008 GTGAACCCCATCTCTCAACAAAAATACAATCACTGGCCGTGGTGGTGGTGGT 10949
 QY 1647 TAGTCCCA--ACGAGAGAGTGGAG--GGGAAATTTGTTGAACCCAGAGAGTGGT 1701
 DB 10948 TAATCCCACTCTCAAGAGCTGAGGTTTGAAGAAATCCGGAACCTGGAGTGGAGGT 10889
 QY 1702 TTGCAGT----GAGATTGAGCACTGCAATCCAGCTGGGCGAGAGAGTGTCTTC 1756
 DB 10888 TGCAGTGAAGCAGAGATGTAACCACTGCACTCAGCTGGTGGTGAAGAACTGCAATC 10829
 QY 1757 AAAAAAAGG 1774
 DB 10828 TCAAGAAAAAAGG 10811

RESULT 7

US-10-109-854-3/c
 ; Sequence 3, Application US/10109854
 ; Patent No. 6650337
 ; GENERAL INFORMATION:
 ; APPLICANT: YAN, Chunhua et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; FILE REFERENCE: C1000758DIV
 ; CURRENT APPLICATION NUMBER: US/10/109,854
 ; PRIOR FILING DATE: 2002-04-01
 ; PRIOR APPLICATION NUMBER: 60/227,470
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: 09/810,671
 ; NUMBER OF SEQ ID NOS: 5
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 21234
 ; TYPE: DNA

ORGANISM: Homo sapiens
US-10-109-854-3

Query Match 5.3%; Score 117.2; DB 4; Length 21234;
Best Local Similarity 76.7%; Pred. No. 3.3e-26;
Matches 198; Conservative 0; Mismatches 48; Indels 12; Gaps 4;

QY 1529 GGGGTGAGAGCAAGTATCATCAGAGGTGAGAGTTGAGACCAACTGGCCCAATG 1588
DB 11068 GGAAGCCGAGGAGGAGAGATCCTGAGGTGAGAGTTGAGACCAACTGGCCCAATG 11009
QY 1589 ATGAACCCCATCTCTACCAAAAAATACAAATCAGCTGCGCTGCTGTGTTG--CCTG 1646
DB 11008 GTGAACCCCATCTCTCTAAAACTATAAAAAATTAGCCGCGGTGTGTTGCTGGCCACTG 10949
QY 1647 TAGTCCCA--AGCAGAGAGGTGAG--GGGAGAAATGCTTGAACCCAGAGGTGTGT 1701
DB 10948 TAATCCCAAGTACTCAGAGAGGCTGAGTGAAGAAATGCGCGCAACTG3AGGTGTGAGGT 10889
QY 1702 TGCAGT-----GAGATTGAGCAACTGCAATCCAGCTGCGGCGAGCGAGTGAAGTGTCTC 1756
DB 10888 TGCAGTGAAGCCGAGATGTAACCACTGCACTCAGCTGCGGTGAGCAAAATGAACTCCATC 10829
QY 1757 AAAAAAAAAAAAAAAAAAG 1774
DB 10828 TCAAGAGAAAAAAAAAG 10811

RESULT 8
US-09-435-739-42
Sequence 42; Application US/09435739
Patent No. 6664105
GENERAL INFORMATION:
APPLICANT: Becker, Itis
APPLICANT: Viodevsky, Israel
APPLICANT: Feinstein, Elena
TITLE OF INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE HAVING HEPARANASE ACTIVITY
TITLE OF INVENTION: EXPRESSION OF SAME IN GENETICALLY MODIFIED CELLS
FILE REFERENCE: 00/20454
CURRENT APPLICATION NUMBER: US/09/435.739
CURRENT FILING DATE: 2001-06-05
NUMBER OF SEQ ID NOS: 47
SOFTWARE: PatentIn version 3.0
SEQ ID NO 42
LENGTH: 44848
TYPE: DNA
ORGANISM: Homo sapiens
US-09-435-739-42

Query Match 5.3%; Score 117; DB 4; Length 44848;
Best Local Similarity 73.1%; Pred. No. 6.7e-26;
Matches 179; Conservative 0; Mismatches 60; Indels 6; Gaps 2;
QY 1529 GGGGTGAGAGCAAGTATCATCAGAGGTGAGAGTTGAGACCAACTGGCCCAATG 1588
DB 42887 GGAAGCCGAGGAGGAGATGATCCTGAGGTGAGAGTTGAGACCAACTGGCCCAATG 42946
QY 1589 ATGAACCCCATCTCTACCAAAAAATACAAATCAGCTGCGCTGCTGTGTTGTTA 1648
DB 42947 GTGAACCCCATCTCTCTAAAAATACAAAAATTAGCTGGGTGTGCTGTGATGCCA 43006
QY 1649 GTCCCAAGCAGAGAGGTGAGGAGAAATGCTGAACCCAGAGAGGTGTGAGTGAAGT- 1707
DB 43007 GCTACTTTGGAGGCTG--GGCAGAGAGATTAATTGAAACCCAGAGAGGAGGTTGAGTG 43065
QY 1708 ---GAGATTGAGCAACTGCAATCCAGCTGGCGAGCGAGTGAAGTGTCTCAAAAAAA 1763
DB 43066 AGCCGAGATGTGCTCACTGCACTCCAGCTGGGTGAGAGAGATTCATCTCAAAAA 43125
QY 1764 AAAAA 1768
DB 43126 AACAA 43130

RESULT 9
US-09-345-882-1/C
Sequence 1; Application US/09345882
Patent No. 639373
GENERAL INFORMATION:
APPLICANT: Bouguetel, Lydie
TITLE OF INVENTION: A NUCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
TITLE OF INVENTION: AND POLYMORPHIC MARKERS ASSOCIATED WITH SAID NUCLEIC ACID.
FILE REFERENCE: GENSER.031A
CURRENT APPLICATION NUMBER: US/09/345,882
CURRENT FILING DATE: 1999-06-30
PRIOR APPLICATION NUMBER: US 60/091,315
PRIOR FILING DATE: 1998-06-30
PRIOR APPLICATION NUMBER: US 60/111,909
PRIOR FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 140
SOFTWARE: Patent.pm
SEQ ID NO 1
LENGTH: 162450
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 72794
OTHER INFORMATION: 5-124-273 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 88073
OTHER INFORMATION: 5-127-261 : polymorphic base A or C
FEATURE:
NAME/KEY: allele
LOCATION: 93714
OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
FEATURE:
NAME/KEY: allele
LOCATION: 97152
OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
FEATURE:
NAME/KEY: allele
LOCATION: 99098
OTHER INFORMATION: 5-130-257 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 99117
OTHER INFORMATION: 5-130-276 : polymorphic base A or G
FEATURE:
NAME/KEY: allele
LOCATION: 103806
OTHER INFORMATION: 5-131-395 : polymorphic base A or T
FEATURE:
NAME/KEY: allele
LOCATION: 106940
OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108106
OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
FEATURE:
NAME/KEY: allele
LOCATION: 108149
OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTT
FEATURE:
NAME/KEY: allele
LOCATION: 108308

```

1 FEATURE:
2 NAME/KEY: allele
3 LOCATION: 97130..97177
4 OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID33
5
6 FEATURE:
7 NAME/KEY: allele
8 LOCATION: 97130..97177
9 OTHER INFORMATION: polymorphic fragment 5-129-144 SEQ ID54
10
11 FEATURE:
12 NAME/KEY: allele
13 LOCATION: 99075..99121
14 OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID34
15
16 FEATURE:
17 NAME/KEY: allele
18 LOCATION: 99075..99121
19 OTHER INFORMATION: polymorphic fragment 5-130-257 SEQ ID55
20
21 FEATURE:
22 NAME/KEY: allele
23 LOCATION: 99094..99140
24 OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID35
25
26 FEATURE:
27 NAME/KEY: allele
28 LOCATION: 99094..99140
29 OTHER INFORMATION: polymorphic fragment 5-130-276 SEQ ID56
30
31 FEATURE:
32 NAME/KEY: allele
33 LOCATION: 103783..103828
34 OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID36
35
36 FEATURE:
37 NAME/KEY: allele
38 LOCATION: 103783..103828
39 OTHER INFORMATION: polymorphic fragment 5-131-395 SEQ ID57
40
41 FEATURE:
42 NAME/KEY: allele
43 LOCATION: 106918..106966
44 OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID37
45
46 FEATURE:
47 NAME/KEY: allele
48 LOCATION: 106918..106966
49 OTHER INFORMATION: polymorphic fragment 5-133-375 SEQ ID58
50
51 FEATURE:
52 NAME/KEY: allele
53 LOCATION: 108084..108130
54 OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID38
55
56 FEATURE:
57 NAME/KEY: allele
58 LOCATION: 108084..108130
59 OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
60
61 FEATURE:
62 NAME/KEY: allele
63 LOCATION: 108127..108177
64 OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
65
66 FEATURE:
67 NAME/KEY: allele
68 LOCATION: 108127..108177
69 OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
70
71 FEATURE:

```

	Query Match	5.3%;	Score 117;	DB 4;	Length 162450;
	Best Local Similarity	76.4%;	Pred. No. 1.7e-25;		
	Matches 198;	Conservative 0;	Mismatches 50;	Indels 11;	Gaps 4;
QY	1529	GGGTGTGAAGGCAAGTAAATACAGAGGGTCAGAGGTTGAGACCAACTGGCCAAATG			1588
Db	86346	GGAAGCGAAGCAGAGTGATCACTGAGGTCAAGAGTTTGAGACCAAGCTGGCCAAAAG			86287
QY	1589	ATGAACCCCATCTTACCAAAAAAATCAATGCTGGC-CGTGCTGTGCTCTGT			1647
Db	86286	GTCGAACCTCGTCTTCACTAAAAATCAAAAAATTAGCCAGCATGATGGCGGTGCTGT			86227
QY	1648	AGTCCCA---ACCCAGAGGTTGAGG--GGAGATTGGCTTGAACCCAGAGGTGGTGT			1702
Db	86226	AATCCCAAGTACTCTGGAGGCTGAGCAGAGAAATTGCTGAAACCCGGAGGCGGAAGTT			86167

Qy	1703	GCAGTGG-----ATTGAGCAATCGCAATCCAGCTGGGGGAGAGAGTGAAGTGGCTCA	1757
Db	86166	GCAGTGGCCGATATATCCACCACTGCACTCCAGCCTGGGGAGAGAGTGAAGTCCCTCT	86107
Qy	1758	AAAAAAAAAAAAAAAAAGAT	1776
Db	86106	CAAAAAAAAAAAAAAAAAAGAT	86088

```

RESULT 10
US-09-759-359A-3/C
: Sequence 3, Application US/09759359A
: Patent No. 6492153
: GENERAL INFORMATION:
: APPLICANT: ABU-THREIDEH, Jane et al
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: CLO01043
: CURRENT APPLICATION NUMBER: US/09/759,359A
: CURRENT FILING DATE: 2001-01-16
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 90541
: TYPE: DNA
: ORGANISM: Human
: US-09-759-359A-3

```

Query Match	5.3%;	Score 116.8;	DB 4;	Length 90541;
Best Local Similarity	74.0%;	Pred. No. 1.3e-25;		
Matches 191; Conservative	0;	Mismatches 57;	Indels 10;	Gaps 3;

Qy	1529	GGGGTGTGAGGCAAGTAAATCAAGAGAGGTGAGAGTTTGAGACCAACTGTGGCAACATG	1588
Db	5725	GGAGGCTGAGGCAAGAAATTACTTGTAGGTGAGAGTTCAAGACCAAGCTGGCCAACTG	5666
Qy	1589	ATGAACCCCATCTCTTACCAAAAAAATACAATCAAGCTGGC-CGTGTGTGTGCTCTGT	1647
Db	5665	GTCGAACCCCATCTCTTACTAAAAATATACAAAGTTTACCAAGGCTGTGGTGTGAGCTGT	5606
Qy	1648	AGTCCCA---ACGAGAGAGTTGAGGGAGAAATTGCTTGAACCCAGAGAGTGAGTTGC	1704
Db	5605	AGTTCACGTACTCAGAGAGGCTGGCAAGAGATTGCTTGAACCCAGAGAGTGAGTTGC	5546
Qy	1705	AGTGAAGATTGAG-----CAACTGCAATTCACGCTGGGCGACGAGTGAGACTGTCTCAA	1758
Db	5545	AGTGAAACAGAGATCAACCACTGCACTGCACTCAAGGCTGTGGTGACAGAGGAGACTCCATCTC	5486
Qy	1759	AAAAAAAAAAAAAAAAAGAT	1776
Db	5485	AAAAAAAAAAAAAAAAAT	5468

```

RESULT 11
US-09-918-686-1
: Sequence 1, Application US/09918686
: Patent No. 6475729
: GENERAL INFORMATION:
: APPLICANT: Brunkow, Mary
: APPLICANT: Proff, Sean
: APPLICANT: Paepeler, Bryan
: APPLICANT: Staehling-Hampton, Karen
: TITLE OF INVENTION: METHODS FOR IDENTIFYING
: TITLE OF INVENTION: GENOMIC DELETIONS
: FILE REFERENCE: 240083.515
: CURRENT APPLICATION NUMBER: US/09/918.686
: CURRENT FILING DATE: 2001-07-30
: NUMBER OF SEQ ID NOS: 105
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 92339

```

```

: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: 7043, 8369, 8401
: OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1

```

Query Match	5.3%	Score 116.8;	DB 4;	Length 92239;
Best Local Similarity	77.0%;	Pred. No. 1.3e-25;		
Matches 197; Conservative	0;	Mismatches 47;	Indels 12;	Gaps 4;

Qy	1529	GGGTGTGGAGGCAAGTAATTCACAGAGGTCAGAGATTGAGACCAACTGGCCAAATG	1588
Db	85287	GGAGGCCGAGGCGAGTGGATCACCCTGAGGCCAGAGATTGGAACCAAGCCTGGCCAAATG	85346
Qy	1589	ATGAAACCCCATCTTACCAAAAAAATACAAATCAGCTGGCCGCTGCG--TGTCCTTG	1646
Db	85347	GTGAAACCTGCTCTTACTACAAATTCAAAAAATTGCTGGGTGTGTGTGGCATATCCTG	85408
Qy	1647	TAGTCCCA---ACGACGAGAGTTGAGG--GGAAGATTGCTTGAACCCAGAGAGTGTGTG	1701
Db	85407	TAGTTCCAGTACTCTCGAAGCTGAGGCGAGAAATTTGTTGAATCTCAGAGGCGAAGT	85466
Qy	1702	TGCAGTGAG----ATTGAGCACTGCAATCCAGCCTGGGCGACGAGTGTAGACTGTCTC	1756
Db	85467	TGCAGTGAGCCAACTATGGCCCATTTGCATCTCCAGCCTGGGTGACAGAGGGAATCTGTGC	85528
Qy	1757	AAAAAAAAAAAAAAAAAAAAA	1772
Db	85527	TCAAAAAAAAAAAAAAAAA	85542

```

RESULT 12
US-09-800-960-3/c
: Sequence 3, Application US/09800960
: Patent No. 6387677
:
: GENERAL INFORMATION:
:
: APPLICANT: YE, Jane et al.
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: CL001158
: CURRENT APPLICATION NUMBER: US/09/800,960
: CURRENT FILING DATE: 2001-03-08
:
: NUMBER OF SEQ ID NOS: 4
:
: SOFTWARE: FastSeq for Windows Version 4.0
:
: SEQ ID NO 3
: LENGTH: 62804
:
: TYPE: DNA
:
: ORGANISM: Human
:
: FEATURE:
:
: NAME/KEY: misc_feature
: LOCATION: (1)..(62804)
:
: OTHER INFORMATION: n = A,T,C or G
:
: US-09-800-960-3

```

Query Match	5.2%;	Score 116.2;	DB 4;	Length 62804;
Best Local Similarity	78.2%;	Pred. No. 1.6e-25;		
Matches 179;	Conservative	0;	Mismatches 43;	Indels 7;
			Gaps	3;

QY	1551	CAGAGAGTCAGGAGTTTGTAGACCAACTCTGGCCAAATGATGTAACCCTCATCTCCAAA	1610
Db	29739	CATGAATATCAGGAGATCGAACAACATCTCTAACATCACCGTGTAACCCCATCTCTCTAA	29680
QY	1611	AAAAATATCAAAATCAGCTGGCCGCCTGCTGTG--CCTGTATGCCCCA---AGCAGAAGGT	1665
Db	29679	AATATCAAAAAATTATGCGCGGCGCTGTGTGGCGGCACCTGTATGTCCTCACTCGGAGGC	29620
QY	1666	TGAGG--GGAGATTTGCTTGAACCCACAGAGCGGTGTGTGACGTATGTAGCAACTGC	1723
Db	29619	TGAGGAGAGAGATGTGCATGAACCCGGAGGSCAGAGCTTGCACTATAGAACCGGCCACTGC	29560

Mon Mar 29 09:56:41 2004

us-09-622-964-1_copy_13900_16115.rn1

Page 9

Search completed: March 25, 2004, 07:24:26
Job time : 170 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2004, 05:47:23 ; Search time 1271 Seconds
(without alignments)
6491.101 Million cell updates/sec

Title: US-09-622-964-1_COPY_13900_16115

Perfect score: 2216
Sequence: 1 aactctgtcttcctccagc.....cctaatgagtttaataac 2216

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2458946 seqs, 1861504846 residues

Total number of hits satisfying chosen parameters: 4917892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1631.6	73.6	1717	15	US-10-264-237-259 Sequence 259, App
2	643.8	29.1	1263	10	US-09-746-783-3 Sequence 3, Appl1
3	404.2	18.2	459	10	US-09-918-995-26929 Sequence 26929, A
4	339.4	15.3	1236	13	US-10-071-766-10 Sequence 10, Appl
5	164	7.4	1198	9	US-09-880-107-2174 Sequence 2174, Ap
6	133.2	6.0	747	15	US-10-027-632-173516 Sequence 173516,
7	133.2	6.0	65042	14	US-10-229-124-3 Sequence 3, Appl1
8	133.2	6.0	149480	10	US-09-873-367C-284 Sequence 284, App
9	133.2	5.8	322101	14	US-10-060-902-1 Sequence 285, Appl
10	127.6	5.7	22645	10	US-09-764-891-7673 Sequence 1, Appl1
11	126.2	5.7	22645	10	US-09-764-891-7673 Sequence 7673, Ap
12	126.2	5.7	22645	10	US-09-764-891-7673 Sequence 8183, Ap
13	124.2	5.6	39000	10	US-09-957-956-5 Sequence 5, Appl1
14	124.2	5.6	130320	15	US-10-408-168-1 Sequence 1, Appl1
15	123.8	5.6	14874	10	US-09-764-891-7672 Sequence 7672, Ap

16	123.4	5.6	753	15	US-10-027-632-21017 Sequence 21017, A
17	123.4	5.6	10500	14	US-10-312-495-13 Sequence 13, Appl
18	123.4	5.6	358246	15	US-10-292-798-1095 Sequence 1095, Ap
19	123.2	5.6	52216	9	US-09-747-810-1 Sequence 1, Appl1
20	123	5.6	7720	9	US-09-954-456-946 Sequence 946, App
21	123	5.6	7720	9	US-09-954-456-1589 Sequence 1589, Ap
22	122.8	5.5	433	15	US-10-027-632-256214 Sequence 256214,
23	122.4	5.5	1094	15	US-10-027-632-256287 Sequence 256287,
24	122.2	5.5	23748	10	US-10-027-632-256666 Sequence 256666,
25	121.8	5.5	37818	10	US-09-764-891-7917 Sequence 7917, Ap
26	121.8	5.5	39703	14	US-10-017-161-1041 Sequence 1041, Ap
27	121.8	5.5	39703	15	US-10-292-798-883 Sequence 883, App
28	121.8	5.5	134292	12	US-10-240-425-1102 Sequence 1102, Ap
29	121.6	5.5	36159	13	US-10-135-687-3 Sequence 3, Appl1
30	121.4	5.5	733	15	US-10-027-632-14278 Sequence 14278, A
31	121.4	5.5	733	15	US-10-027-632-14279 Sequence 14279, A
32	121.2	5.5	433	15	US-10-027-632-256213 Sequence 256213,
33	121.2	5.5	11282	9	US-09-764-847-1747 Sequence 1747, Ap
34	121.2	5.5	11282	14	US-10-092-154-1747 Sequence 1747, Ap
35	121.2	5.5	13630	9	US-09-764-868-1369 Sequence 1369, Ap
36	121.2	5.5	13630	11	US-10-017-161-2067 Sequence 2067, Ap
37	121.2	5.5	21221	14	US-10-017-161-2067 Sequence 2067, Ap
38	121	5.5	761	15	US-10-027-632-25603 Sequence 25603, A
39	121	5.5	21721	9	US-09-764-853-861 Sequence 861, App
40	121	5.5	162025	14	US-10-272-665-35 Sequence 35, Appl
41	121	5.5	162025	14	US-10-272-665-36 Sequence 36, Appl
42	121	5.5	162025	14	US-10-273-321-35 Sequence 35, Appl
43	121	5.5	162025	14	US-10-273-321-36 Sequence 36, Appl
44	121	5.5	162025	14	US-10-272-756-35 Sequence 35, Appl
45	121	5.5	162025	14	US-10-272-756-36 Sequence 36, Appl

ALIGNMENTS

RESULT 1
US-10-264-237-259
; Sequence 259, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Btise et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PAL31P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 259
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1072)..(1072)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1596)..(1596)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1634)..(1634)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1657)..(1657)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc_feature

Query Match	Score	DB	Length
Best Local Similarity	73.6%;	1631.6;	1717;
Matches 1692; Conservative	98.5%;	Pred. No. 0;	
	3;	Mismatches	10

Qy	57	GAGAGCGAGAGAGATGCTTCAACGCTGGCATTCATGTGGCGCTTCCATAGGCTCTGCAGTCCAT	116
Db	1	GAGGAGGAGAGAGATGCTTCAACGCTGGCATTCATGTGGCGCTTCCATAGGCTCTGCAGTCCAT	60
Qy	117	GATCACCATCTCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGATCC	176
Db	61	GATCACCATCTCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAGGATCC	120
Qy	177	CTTCTCCAGAGGGCTTGCACCAAAACCAAGAGCGCAAAACAGAACTTATGGGGCAG	236
Db	121	CTTCTCCAGAGGGCTTGCACCAAAACCAAGAGCGCAAAACAGAACTTATGGGGCAG	179
Qy	237	GAAAGCAACAAGGCGCTGAAAGCTTAAGCTGTGACGCTTCAAGTCTGCCCACTGAT	296
Db	180	GAAAGCAACAAGGCGCTGAAAGCTTAAGCTGTGAGCGCTTCAAGTCTGCCCACTGAT	239
Qy	297	CAGAGGCCAGCTTACTACAGTGCCTCCACAGAGCGCCCTGAGCCCACTGCCATGTTCTTC	356
Db	240	CAGAGGCCAGCTTACTACAGTGCCTCCACAGAGCGCCCTGAGCCCACTGCCATGTTCTTC	299
Qy	357	CCCTTAGAACCATCAGGCGCGTCMAAGCTTCACTGTCCAAGCATATGACCAAAAGAC	416
Db	300	CCCTTAGAACCATCAGGCGCGTCMAAGCTTCACTGTCCAAGCATATGACCAAAAGAC	359
Qy	417	AAAAGCTTAAAGCTGTGACTTCTGGGCGCAAGAAAGTTTGAATGTGCTCAGAGAGC	476
Db	360	AAAAGCTTAAAGCTGTGACTTCTGGGCGCAAGAAAGTTTGAATGTGCTCAGAGAGC	419
Qy	477	GATGGGGCTTGTATGAGACCCAGAACTATCTCACTGAGAGAGAAACTGTGAGTT	536
Db	420	GATGGGGCTTGTATGAGACCCAGAAATATCTCACTGAGAGAGAAACTGTGAGTT	479
Qy	537	AACTTAGCGGATATCCAGAGATCCCGGAAATCACTCAAGAACTTTGGAAACAACA	596
Db	480	AACTTAGCGGATATCCAGAGATCCCGGAAATCACTCAAGAACTTTGGAAACAACA	539
Qy	597	CCAAACCAATACACATACACTCAAGATACATGAGATCTTATTTGGCCCTTGGAAAC	656
Db	540	CCAAACCAATACACATACACTCAAGATACATGAGATCTTATTTGGCCCTTGGAAAC	599
Qy	657	AGGTGTGTCTCCACTGAAACAGGGGCACTGCATTTGCCCTGTGCCCAACCCAGCTTC	716
Db	600	AGGTGTGTCTCCACTGAAACAGGGGCACTGCATTTGCCCTGTGCCCAACCCAGCTTC	659
Qy	717	CTTGTCTGAGCTTACCTTCTCTCAACAATTTCTAGGGTTCAATCACTGCAGAGACA	776
Db	660	CTTGTCTGAGCTTACCTTCTCTCAACAATTTCTAGGGTTCAATCACTGCAGAGACA	719
Qy	777	CTGAGACTTAGCCACAGACTGGCTTGGGGTATATCTTGGCACACTTCAAGGAGATCTTA	836
Db	720	CTGAGACTTAGCCACAGACTGGCTTGGGGTATATCTTGGCACACTTCAAGGAGATCTTA	779
Qy	837	GGGAAAGTTCGGGACCTTTTCTCACTTCACTCGTGTATCAACCGGAAAGCTTTTGGGA	896
Db	780	GGGAAAGTTCGGGACCTTTTCTCACTTCACTCGTGTATCAACCGGAAAGCTTTTGGGA	839
Qy	897	CCAGGTGAAGAAATGAGTTGTGCTGACCAAGATCTCTGAGAACCTGCCAGGGC	956
Db	840	CCAGGTGAAGAAATGAGTTGTGCTGACCAAGATCTCTGAGAACCTGCCAGGGC	899
Qy	957	TGACAGGCCAGGCTTAGCTGAGCAGATGTATCACTGGCCCCCAACTTACTTTGAGCAAGG	1016

D	b		900	TGACAGCCAGGCTTAGCTGAGCAATGTATTCACTGGCCCCAACCTTACTTTGACGAAG	959
O	y		1017	GTCGCTGACCACAAACCATGAGTGCGAGTCAGCAGTAACAATTAACAATTCCCCA	107
D	b		960	GTGGCTGACCCAAAACCATGAGTGGCAGTCACTGTGATGACGATTAACATCTTCCCCA	101
O	y		1077	TAACATTTAAGGCTGTATGCCAACATACAGGAAAGGTGGCAGAACTGCCTCACTCC	113
D	b		1020	TAACTATTTAAGGTATGTAACCCAAACACTTACAGGAAAGGTGGCAGAACT-CTNCACTCC	107
O	y		1137	TAGGAACGTGTATGATGTGAGGTTGAGGGTGTCAAGGCCCTTAAGTCAATTTTTCAC TG	119
D	b		1079	TAGGAACGTGTATGATGTGAGGTTGAGGGTGTCAAGGCCCTTAAGTCAATTTTTCAC TG	1138
O	y		1197	CCTGGAACTCACCAAATACTTCTTGCTTCTTGGGGTCAGCCCAAGCTGTACAAA	1256
D	b		1139	CTGGGAACCTCACCAAATACTTCTTGCTTCTTGGGGTCAGCCCAAGCTGTACAAA	1198
O	y		1257	ATCAGATATTTCCCTTTATTCAGATTTCTGTGACACTGTCAACCAATATTAACA CCCC	1316
D	b		1199	ATCAGATATTTCCCTTTATTCAGATTTCTGTGACACTGTCAACCAATATTAACA CCCC	1258
O	y		1317	ACTTCAGCCCCCATCACTGCTGGAGGAGAAGTAACTTCCCTTTTGTGATTCYCAAGCAGT	1376
D	b		1259	ACTTCAGCCCCCATCACTGCTGGAGGAGAAGTAACTTCCCTTTTGTGATTCYCAAGCAGT	1318
O	y		1377	TACTTCAAGGGCTGAGAAACAGCAGCTATTAATGATTGAACCTTAAAAAGGCAACA TTT	1436
D	b		1319	TACTTCAAC-GGTCAAGAACAGCAGCTATTAATGATTGAACCTTAAAAAGGCAACA TTT	1377
O	y		1437	CANTCTTGCTTTAGGCTTAAGA CAGAACTTGGCAAA CATCTGGGCTGTCTAGCAAA G	1496
D	b		1378	CAGTCTTGCTTTAGGCTTAAGA CAGRACTTGGCAAA CATCTGGGCTGTCTAGCAAA G	1437
O	y		1497	GATGTTCAATTTAAGATCTTGTCTTGCGCTGGCTGTGGAGGCAAGTAA TCACAGAG	1556
D	b		1438	GATGTTAATTTAAGATCTTGTCTTGCGCTGGCTGTGGAGGCAAGTAA TCACAGAG	1497
O	y		1557	GTCAGAGATTTGAACCAACCTGGCCAACATGATGAAACCCCATCTTACCAAAAAA AT	1616
D	b		1498	GTCAGAGATTTGAACCAACCTGGCCAACATGATGAAACCCCATCTTACCAAAAAA AT	1557
O	y		1617	ACAAATCAGCTGCGCTGTGAGTGTGCTGTAGTCCCAACGACAGAGTTGAGGGAGAA	1676
D	b		1558	ACAAATCAGCTGCGCTGTGAGTGTGCTGTAGTCCCAACGACAGAGTTGAGGGAGAA	1617
O	y		1677	TTGCTTAACCCAGAGAGTGTGTGACGTAGA TTGA-GCAATGTCAATCCAGCCTGG	1735
D	b		1618	TTGCTTAACCCAGAGAGTGTGTGACGTAGA TTGANCAACTGSAATCCAGCCTGG	1677
O	y		1736	GCAGC-AGAGTGAAGTGTCAAAAAA AAAAAAAAAA	1772
D	b		1678	GCAGCAGGAGTGAAGTGTCAAAAAA AAAAAAAAAA	1715

RESULT 2
 US-09-746-783-3
 ; Sequence 3, Application US/09746783
 ; Publication No. US20030044935A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacob, Kenneth
 ; McCoy, John M.
 ; LaVallie, Edward R.
 ; Racie, Lisa A.
 ; Treacy, Maurice
 ; Spaulding, Vikki
 ; Agostino, Michael J.
 ; Howes, Steven H.
 ; Fechtel, Kim
 ;
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
 ;
 ; NUMBER OF SEQUENCES: 231
 ;

CORRESPONDENCE ADDRESS:
ADDRESS: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/746,783
FILING DATE: 21-Dec-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Milasincic, Debra J.
REGISTRATION NUMBER: 46,931
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1263 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-746-783-3

Query Match 29.1%; Score 643.8; DB 10; Length 1263;
Best Local Similarity 99.7%; Pred. No. 1,7e-196;
Matches 645; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

17 CAGCCGGAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 76
225 CAGCCGGAACAAAGAGAGATGAGTTCAGCCCAATCAGAGAGAGAGATGCTCA 284
77 CGGTGATCATTTGGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 136
285 CGGTGATCATTTGGCGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 344
137 AAACTCAAGAGACCAATCTAGTGTGCGCCAAAGAGAGATCCCTTCTTCTTCTT 196
345 AAACTCAAGAGACCAATCTAGTGTGCGCCAAAGAGAGATCCCTTCTTCTTCTT 404
197 CAAAAACCAAG 256
405 CAAAAACCAAG 464
257 GCTTAAAGGCTGTGAGAGCGCTTCAAGTCTGCGCCCACTGTATCAGAGGCGCTT 316
465 GCTTAAAGGCTGTGAGAGCGCTTCAAGTCTGCGCCCACTGTATCAGAGGCGCTT 524
317 TGGCCCAAGAGAGCGCGCTTCAAGTCTGCGCCCACTGTATCAGAGGCGCTT 376
525 TGGCCCAAGAGAGCGCGCTTCAAGTCTGCGCCCACTGTATCAGAGGCGCTT 584
377 GTCAAGAGCTTCAAGTGTCAAGGAGTACCAAGAGAGAGAGAGAGAGAGAGAGAG 436
585 GTCAAGAGCTTCAAGTGTCAAGGAGTACCAAGAGAGAGAGAGAGAGAGAGAGAG 644
437 TTCTGGGGGCAAG 496
645 TTCTGGGGGCAAG 704
497 CCCAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 556
705 CCCAGAGATATCTCAAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 764
557 GATCCCGGAGAGATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 616

Db 765 GATCCCGGAGAGATCACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 824
QY 617 ACTCAAGATACATGATGATCTTATTTGGGCGCTTGAAGAGAGAGAGAGAG 663
Db 825 ACTCAAGATACATGATGATCTTATTTGGGCGCTTGAAGAGAGAGAGAGAG 871

RESULT 3
US-09-918-995-26929
Sequence 26929, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hsueh, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918,995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235,076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for windows Version 3.0
SEQ ID NO 26929
LENGTH: 459
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(459)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-26929

Query Match 18.2%; Score 404.2; DB 10; Length 459;
Best Local Similarity 96.4%; Pred. No. 2.3e-119;
Matches 423; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 1333 CGTGGAGAGAGATGATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1392
Db 22 CGAGGAGATGATGATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 80
QY 1393 AACACGACCTATTTATGATTTGAAGCTTAAAGAGAGAGAGAGAGAGAGAGAG 1452
Db 81 AACACGACCTATTTATGATTTGAAGCTTAAAGAGAGAGAGAGAGAGAGAGAG 140
QY 1453 CTAAAGACAGAACTTGGCAAACTGTGCGCTTCTTCTTCTTCTTCTTCTTCT 1512
Db 141 GCTAAGACAGAACTTGGCAAACTGTGCGCTTCTTCTTCTTCTTCTTCTTCT 200
QY 1513 AATCTTGTCTTGGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1572
Db 201 AATCTTGTCTTGGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 260
QY 1573 CAACCTGGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1632
Db 261 CAACCTGGCCCAATGATGATGATGATGATGATGATGATGATGATGATGATG 320
QY 1633 TGT 1692
Db 321 TGT 380
QY 1693 GGT 1752
Db 381 GGT 440
QY 1753 TCTCAAAAAAAAAAAAAA 1771
Db 441 TCTCAAAAAAAAAAAAAA 459

RESULT 4
US-10-071-766-10/c
Sequence 10, Application US/10071766

Publication No. US20020192678A1
GENERAL INFORMATION:
APPLICANT: Huel-Mei Chen
TITLE OF INVENTION: GENES EXPRESSED IN SENESENCE
FILE REFERENCE: PA-0043 US
CURRENT APPLICATION NUMBER: US/10/071,766
CURRENT FILING DATE: 2002-02-07
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PERL Program
SEQ ID NO 10
LENGTH: 1326
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20020192678A1 1138151.2
US-10-071-766-10

Query Match 15.3%; Score 339.4; DB 13; Length 1326;
Best Local Similarity 99.7%; Pred. No. 4.2e-98;
Matches 340; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1876 AGCAGATTCCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGGAGGCTCTCACTGTGT 1935
DB 1332 AGCAGATTCCTTAACCTGCTTCTTAATGGGATGCTTGGCCAGGAGGCTCTCACTGTGT 1267
QY 1936 GTACACGACGAGACACTGATCCGTCACAGCCATACAGCTGTCCACACTGGAAGACGTG 1995
DB 1266 GTACACGACGAGACACTGATCCGTCACAGCCATACAGCTGTCCACACTGGAAGACGTG 1207
QY 1996 TCCACAGACGCTGTAATCAATGTTAGTTATGATTAATAATCCAGACTTATAG 2055
DB 1206 TCCACAGACGCTGTAATCAATGTTAGTTATGATTAATAATCCAGACTTATAG 1147
QY 2056 CTTTAATGCTTTTATTAATAAAGCTGGAAGCTGACCTGGAACCACTTGAACATTT 2115
DB 1146 CTTTAATGCTTTTATTAATAAAGCTGGAAGCTGACCTGGAACCACTTGAACATTT 1087
QY 2116 AACTCAGACTCTGATTCAGAGTGGGAAACCTTATGTTATGATTAATAATCCAGACCA 2175
DB 1086 AACTCAGACTCTGATTCAGAGTGGGAAACCTTATGTTATGATTAATAATCCAGACCA 1027
QY 2176 CACCTTAGTATCTGCCCAACTAATGAGTTAATAATAC 2216
DB 1026 CACCTTAGTATCTGCCCAACTAATGAGTTAATAATAC 986

RESULT 5

US-09-880-107-2174/c
Sequence 2174, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054
NUMBER OF SEQ ID NOS: 3950
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2174
LENGTH: 1198
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Genbank Accession No. US20020142981A1 L20941
US-09-880-107-2174

Query Match 7.4%; Score 164; DB 9; Length 1198;
Best Local Similarity 100.0%; Pred. No. 1.9e-41;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2053 CAGCCTTAAATGCTTTATTAATCAATTAAGTGAAGCTGACATGACATTTGAACA 2112
DB 1198 CAGCCTTAAATGCTTTATTAATCAATTAAGTGAAGCTGACATGACATTTGAACA 1139
QY 2113 TTTAAGTGAAGCTGATTCAGAGTGGGAAACCTTATTTATCTGAATCCAGACAG 2172
DB 1138 TTTAAGTGAAGCTGATTCAGAGTGGGAAACCTTATTTATCTGAATCCAGACAG 1079
QY 2173 CCACACTTATGATTAATCTGCCCAACTAATGATTTAATAATAC 2216
DB 1078 CCACACTTATGATTAATCTGCCCAACTAATGATTTAATAATAC 1035

RESULT 6

US-10-027-632-173516/c
Sequence 173516, Application US/10027632
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 173516
LENGTH: 747
TYPE: DNA
ORGANISM: Human
US-10-027-632-173516

Query Match 6.0%; Score 133.2; DB 15; Length 747;
Best Local Similarity 82.8%; Pred. No. 1.2e-31;
Matches 202; Conservative 0; Mismatches 33; Indels 9; Gaps 4;

QY 1537 AGCAGATGAATCAAGAGGTGAGAGTTGAGACCACTGGCCAAATATATAAACC 1596
DB 454 AGCAGATGAATCAAGAGGTGAGAGTTGAGACCACTGGCCAAATATATAAACC 395
QY 1597 CCATCTCAACCAAAAAATTAACATCACTGCGCTGTG--GTGTGCTGTAGTCCCA 1654
DB 394 CCGTCTCTA-CTAAAAATTAACATCACTGCGCTGTGCGTGTGCGCTGTATCCCA 336
QY 1655 ---AGCAGAGGTGAG--GGAGAAATTTGCTGAACCCAGAGGTGTGTGCAAGT 1708
DB 335 GCTACTCAGAGGTGAGGCAAGGAAATTTGCTGAACCCAGAGGTGTGTGCAAGT 276
QY 1709 AGATTGAGCACTGCAATCACTGCGCGCAAGAGTGAAGTCTCTCAAAAAA 1768
DB 275 AGATTGAGCACTGCACTCACTGCGCGTGAAGAGTGAAGTCTCACTCAAAAAA 216
QY 1769 AAAA 1772
|||||


```

P R I O R F I L I N G D A T E : 2000-11-01
P R I O R A P P L I C A T I O N N U M B E R : U . S .   60 / 245 , 084
P R I O R F I L I N G D A T E : 2000-11-01
N U M B E R O F S E Q I D N O S : 1067
S O F T W A R E : PatentIn version 3.0
S E Q I D N O 284
L E N G T H : 149480
T Y P E : DNA
O R G A N I S M : Homo sapiens
US-03-873-367C-284

Query Matchch          6.0%; Score 133.2; DB 10; Length 149480;
Best Local Similarity  78.5%; Pred. No. 4,7e-30;
Matches 201; Conservative 0; Mismatches 43; Indels 12; Gaps 3;

Qy      1529  GCGGTGAGGACCAAGTCATCACAGAGGTCAGAGTTTGAGACCACCTGGCCAAATG 1588
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      8803  GGAGGCTGAAGCAGGTGATCATCTTGAGGTTCAGAGTTTAAGACCAAGCTGGCCAAATG 8862

Qy      1589  ATGAACCCCATCTCTACCAAAAAAATAACAATACAGCTGGCGTCTGTGTG--TGCGTG 1646
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      8863  GTAAAACCCCATCTCTACCAAAAAAATAAAAAAATTAGCTGGGTGTATGGTGATGCCTG 8922

Qy      1647  TAGTCCCAACGACGAGAG-----TTGAGGGGAGAATTGCTTTGAACCAAGAGGTGTGCT 1701
           ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      8923  TAGTCCCAAGTACTGGGAGGCTGAGACAGAGAAATTGCTTGATTCAGAGAGGTGAGGT 8982

```

```

OY      1702 TGCAGT-----GAGATTGGACACTGTCCATTCAGCCTCGGCGCACGGAATGAGACTGTCTTC 1756
                                     |||          |||          |||          |||          |||
Db      8983 TGCAGTAAAGCCGAGATGCTGCCACTTGCACTCCAGCCTGGGTGACAGAGCATCTGTC   9042

OY      1757 AAAAAAAAAAAAAA    1772
                                     |||          |||          |||          |||
Db      9043 TCAAAAAAAAAAAAAAA    9058


RESULT 9
US-09-873-367C-285
; Sequence 285, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; APPLICANT: Soppet, Daniel
; APPLICANT: Endress, Gregory
; APPLICANT: Augustus, Meena
; APPLICANT: Edner, Reinhard
; APPLICANT: Carter, Kenneth
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
; FILE REFERENCE: 689290-64
```

```

? PRIOR APPLICATION NUMBER: U.S. 60/236,891
? PRIOR FILING DATE: 2000-09-29
? PRIOR APPLICATION NUMBER: U.S. 60/236,842
? PRIOR FILING DATE: 2000-09-29
? PRIOR APPLICATION NUMBER: U.S. 60/244,867
? PRIOR FILING DATE: 2000-11-01
? PRIOR APPLICATION NUMBER: U.S. 60/245,084
? PRIOR FILING DATE: 2000-11-01
? NUMBER OF SEQ ID NOS: 1067
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 285
? LENGTH: 149480
? TYPE: DNA
? ORGANISM: Homo sapiens
US-09-873-367C-285

Query Match      6.0%; Score 133.2; Length 149480;
Best Local Similarity 78.5%; Pred. No. 4.7e-30;
Matches 201; Conservative 0; Mismatches 43; Indels 12; Gaps 3;

```

US-09-764-891-7673/C
 : Sequence 7673, Application US/09764891
 : Publication No. US20030077808A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 : FILE REFERENCE: PC006
 : CURRENT APPLICATION NUMBER: US/09/764,891

Query Match	5.7%	Score 26.2	DB 10	Length 22645
Best Local Similarity	78.8%	Pred. No. 2.3e-26		
Matches	204	Conservative	0	Mismatches 43; Indels 12; Gaps 4
1529	GGGTGTGAGCGCAAGTAATCACAGAGGTCAGGAGTTTGAGCAACCTGGCCCAATG	1588		
11375	GGAAGCTGAGGCGAGCGGATCACTTGGAGTTCAGGAGTTTCAGAGATTCAGCCATG	11316		
1589	ATGAACCCCATCTCTTCACAAAAAATTAACAATCACTGGCGCTCGTGCG--TGCTCG	1646		
11315	GTCAAACCTCTGTCTCTACCAAAAAATACAAAATTAATTGATGTGTGGTGCATGCTTG	11256		
1547	TAGTCCCA---ACGACAGAGTTGAG--GGAGAAATTGCTTGAACCCAGAGGCGTGCT	1701		
11255	TAATCCAGACTCTTGGAGGCTGAGGTTGGAGATCACTTGAACCCAGAGGTCGAGGT	11196		
1702	TGCAGT-----GATTGAGCACTCGAATACAGACCTGGCGCGACGAGTGAAGTGTCTC	1756		
11195	TGCATAGCTGAAGATTGTGTCATTGACATTCACGCTGGGGCAACAGAGGGAGACTGTTTC	11136		

QY 1757 AAAAAAAAAAAAAAGA 1775
11135 AAAAAAAAAAAAAAGA 11117

RESULT 13

US-09-957-956-5
; Sequence 5, Application US/09957956
; Publication No. US20030130215A1
; GENERAL INFORMATION:
; APPLICANT: Ryan, James W.
; TITLE OF INVENTION: Genomic Polynucleotide Fragments From Chromosome 7
; FILE REFERENCE: US-14,000-US
; CURRENT APPLICATION NUMBER: US/09/957,956
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/234,422
; PRIOR FILING DATE: 2001-09-21
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 39000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-957-956-5

Query Match 5.6%; Score 124.2; DB 10; Length 39000;
Best Local Similarity 76.3%; Pred. No. 1.5e-27;

Matches 196; Conservative 0; Mismatches 48; Indels 13; Gaps 3;

QY 1529 GGGGTGTGAGGCAAGTATCATCAGAGGTTCAGAGCTTGACCAATCGGCCAATG 1588
Db 36491 GGAGGCTGAGGCGAGTATGATCCTGAGGTCAGAGTTGAAGACCACTGGCCAAACG 36550
QY 1589 ATGAACCCCATCTCTACCAAAAAATACAAATCAGCTGCGCTGCTG--GTGCTG 1646
Db 36551 GTGAACCCCATCTCTCTAATAAAATACAAAAATTAATGCTGGCGTGGTGGGCTG 36610
QY 1647 TAGTCCCAAC-----GCAGAGGTTGAGGGAGAAATTTGCTGAACCCGAGAGTGTG 1700
Db 36611 TAATCCAGTACTCGGAGGCTGAGGCGAGGAAATTTGCTGAACCCGAGAGCGGAGG 36670
QY 1701 TTGCAGTGAAG--ATTGAGCAATCTCAATCCAGCTGGGCGAGGAGTGAAGTGTCT 1755
Db 36671 TTGCAGTGAAGCAAAATCAACACCACTCCAGCTGGGCTGAGAGGAGATTCAT 36730
QY 1756 CAAAAAAGAAAAA 1772
Db 36731 CTCAAAAAAGAAAAA 36747

RESULT 14

US-10-408-168-1/c
; Sequence 1, Application US/10408168
; Publication No. US20030235847A1
; GENERAL INFORMATION:
; APPLICANT: Paepker, Bryan W.
; APPLICANT: Proli, Sean
; APPLICANT: Charney, Patrick R.
; APPLICANT: Brunow, Mary E.
; APPLICANT: Ullertlinden, Andreas Gerardus
; TITLE OF INVENTION: ASSOCIATION OF POLYMORPHISMS IN THE SOST
; TITLE OF INVENTION: ASSOCIATION OF POLYMORPHISMS IN THE SOST
; FILE REFERENCE: 240083.525
; CURRENT APPLICATION NUMBER: US/10/408,168
; CURRENT FILING DATE: 2003-04-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 130320
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

; NAME/KEY: misc feature
; LOCATION: 100752, 100754, 102080, 117731, 124408, 124532, 124585, 124955,
; LOCATION: 124956, 124963, 124964, 124965, 124967
; OTHER INFORMATION: n = A,T,C or G
US-10-408-168-1

Query Match 5.6%; Score 124.2; DB 15; Length 130320;
Best Local Similarity 72.6%; Pred. No. 3.4e-27;

Matches 204; Conservative 0; Mismatches 68; Indels 9; Gaps 3;

QY 1529 GGGGTGTGAGGCAAGTATCATCAGAGGTTCAGAGCTTGACCAATCGGCCAATG 1588
Db 127635 GGAGGCGGAGCGGCGGAGTCACTGAGGTCAAGAGTTTGAACCACTGGCCAAATG 127576
QY 1589 ATGAACCCCATCTCTACCAAAAAATACAAATCAGCTGCGCTGCTG--TGC 1643
Db 127575 GTGAACCCCATCTCTAATAAAATACAAAAATTAATGCTGGCGTGGTGGCGCAGC 127516
QY 1644 CTGTAGTCC--CAACGAGAGGTTGAG--GGAGAAATGCTTGAACCCGAGAGTGTG 1699
Db 127515 CAGTAGCCCAAGCTTCTTGGAGGCTAAGGAGAGAAATTCCTGAACCCGAGAGTGTG 127456
QY 1700 GTTCAAGTGAATGAGCAATCGCAATCCAGCTGGCGAGGAGTGAAGTCTTCAAA 1759
Db 127455 GTTACAGTGAATGAGCGGCACTGCACTCCAGCTGGCGGAGAGAGAGTCCGTCTCA 127396
QY 1760 AAAAAAAAAAAGATGCTCTCACTTGGCTCTCTACT 1800
Db 127395 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 127355

RESULT 15

US-09-764-891-7672/c
; Sequence 7672, Application US/09764891
; Publication No. US2003077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7672
; LENGTH: 14874
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-7672

Query Match 5.6%; Score 123.8; DB 10; Length 14874;
Best Local Similarity 78.8%; Pred. No. 1e-27;

Matches 201; Conservative 0; Mismatches 42; Indels 12; Gaps 4;

QY 1529 GGGGTGTGAGGCAAGTATCATCAGAGGTTCAGAGCTTGACCAATCGGCCAATG 1588
Db 3589 GGAGGCTGAGGCGAGGAGTCACTTGAAGTCAAGAGTCAAGCTGGCCAAACATG 3530
QY 1589 ATGAACCCCATCTCTACCAAAAAATACAAATCAGCTGCGCTGCTG--TGCTG 1646
Db 3529 GTGAACCCCATCTCTACCAAAAAATACAAAAATTAATGCTGGCGTGGTGGCGCAGC 3470
QY 1647 TAGTCCCA--ACGAGAGGTTGAG--GGAGAAATGCTTGAACCCGAGAGTGTGCT 1701
Db 3469 TAATCCAGTACTTGGAGGCTGAGGTGAGAAATCACTTGAACCCGAGAGTGTGAGGT 3410
QY 1702 TGCAGT-----AGATTGAGCAATCGCAATCCAGCTGGGCGAGGAGTGAAGTGTCT 1756
Db 3409 TGCATGAGTGAATGATTTGCTTGCATCTCCAGCTGGGCGAACAAGGAGAGTGTTC 3350
QY 1757 AAAAAAAAAAAAAA 1771
Db 3349 AAAAAAAAAAAAAA 3335

Mon Mar 29 09:56:42 2004

us-09-622-964-1_copy_13900_16115.rnpb

Page 8

Search completed: March 25, 2004, 09:52:04
Job time : 1276 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using bw model

Run on: March 25, 2004, 02:59:44 ; Search time 5671 Seconds
(without alignments)
11668.946 Million cell updates/sec

Title: US-09-622-964-1_COPY_13900_16115

Perfect score: 2216
Sequence: 1 accctctgtctcttcacgc.....ctaatgagtttaataaac 2216

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_estum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estcom:*
17: em_gse_hum:*
18: em_gse_inv:*
19: em_gse_pln:*
20: em_gse_vic:*
21: em_gse_fun:*
22: em_gse_man:*
23: em_gse_mus:*
24: em_gse_pro:*
25: em_gse_rnd:*
26: em_gse_pbg:*
27: em_gse_vrl:*
28: gb_gse1:*
29: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	626.2	28.3	669	12	BQ028631
2	619.2	27.9	668	14	BQ028631 UI-1-EEO-CA441116
3	611.2	27.6	654	13	BU629816
4	611.2	27.6	669	14	CA308356

5	603.4	27.2	773	10	BE410951	BE410951 601303662
C	600.6	27.1	629	14	CA397980	CA397980 C988C05.x
7	581.2	26.2	632	12	BC950642	BC950642 MR1-CT073
8	581	26.2	593	10	BE385296	BE385296 601277572
9	555.8	25.1	636	14	CA397981	CA397981 C988C05.y
10	545	24.6	545	14	CA395098	CA395098 C860C05.y
11	530.6	23.9	545	13	BU657324	BU657324 C122H04.z
12	501.8	22.6	711	13	BK095540	BK095540 BX095540
13	496.6	22.4	523	12	BM97639	BM97639 UI-E-DXO-
14	492.6	22.2	529	12	BM707813	BM707813 UI-E-C11-
15	466.4	21.0	469	9	AA307119	AA307119 EST178031
16	457.4	20.6	479	13	BQ185946	BQ185946 UI-E-EJ1-
C	457.4	20.6	484	13	BQ184248	BQ184248 UI-E-EJ1-
17	453.2	20.5	651	13	BU731149	BU731149 UI-E-C11-
C	430.8	19.4	601	12	BG951790	BG951790 MR1-CT073
C	429.8	19.4	466	13	BQ187559	BQ187559 UI-E-EJ1-
20	429	19.4	484	14	N31453	N31453 yx55d08.r1
21	422	19.0	422	9	A1917923	A1917923 t214G04.x
22	414.4	18.7	421	9	A1074304	A1074304 c256b11.x
C	413.2	18.6	444	14	N33227	N33227 y07d02.g1
23	411.6	18.6	484	14	M41943	M41943 y07d02.r1
25	396.2	17.9	766	10	BE275846	BE275846 601121656
26	362.4	16.4	434	12	BI031472	BI031472 PMO-MT043
27	362.4	16.3	432	12	BI031468	BI031468 PMO-MT043
28	349.2	15.8	486	13	C18742	C18742 C18742 Huma
29	347.4	15.7	474	9	A1190190	A1190190 q437C03.x
C	347.4	15.7	592	12	BM633028	BM633028 UI-E-C10-
C	345.8	15.6	484	13	BU741783	BU741783 UI-E-EJ0-
C	345.8	15.6	659	13	BU730894	BU730894 UI-E-C11-
33	345.4	15.6	537	12	BM691456	BM691456 UI-E-C11-
34	345.4	15.5	537	12	BM691456	BM691456 UI-E-C11-
C	344.2	15.5	674	13	BU731809	BU731809 UI-E-EJ1-
C	342.6	15.5	676	13	BU741926	BU741926 UI-E-EJ1-
36	342.4	15.5	371	12	BM694530	BM694530 UI-E-C11-
37	342.4	15.5	429	12	BI026143	BI026143 CM4-MT036
C	332.2	15.0	479	12	BI480625	BI480625 H2RP-001
38	327.2	14.8	599	12	BG951382	BG951382 MR1-CT073
40	325	14.7	519	12	BM718146	BM718146 UI-E-EJ0-
41	310.4	14.0	419	12	BM685122	BM685122 UI-E-EJ1-
C	298.8	13.5	503	12	BM932117	BM932117 UI-E-EJ1-
42	298.8	13.5	657	29	CE720218	CE720218 t19t-g88-
43	295.4	13.3	492	13	BU726009	BU726009 UI-E-C10-
44	294.6	13.3				
45						

ALIGNMENTS

RESULT 1
LOCUS BQ028631 669 bp mRNA linear EST 27-MAR-2002
DEFINITION UI-1-EEO-ayz-f-11-0-UI.g1 NCI CGAP P17 Homo sapiens cDNA clone
UI-1-EEO-ayz-f-11-0-UI 3', mRNA sequence.

ACCESSION BQ028631
VERSION BQ028631.1 GI:19763910
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATTC

CDNA library preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this CDNA
sequence: 393-621, >Alu

Seq primer: M13 FORWARD
POLYA=Yes

FEATURES
source

Location/Qualifiers
1..669
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-1-EE0-ayz-f-11-0-UI"
/tissue_type="Placenta Chorioncarcinoma"
/lab_host="DH10B (life technologies)"
/clone_lib="NCI CGAP p17"
/note="Organ: Placenta; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP p17 is a cDNA library containing the following tissue(s): Placenta Chorioncarcinoma UAR Line (ATCC). The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is AGGCATCACA.
TAG_TISSUE=human chorioncarcinoma
TAG_LIB=UI-1-EE0
TAG_SEQ=AGGCATCACA"

ORIGIN

Query Match 28.3%; Score 626.2; DB 12; Length 669;
Best Local Similarity 98.2%; Pred. No. 4.5e-138;
Matches 642; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
1159 TTGAGGGTGTCCAGCGCCCTTAGGTATTTCTCACTGCTGGGAACCTCAACCAATAC 1218
17 TTGAGGGTGTCCAGCGCCCTTAGGTATTTCTCACTGCTGGGAACCTCAACCAATAC 1218
1219 TTCTTCCTTCCTTGGGGGTGAGCCCAAGGTGCACAAATACAGATATTTCTTATTC 1278
77 TTCTTCCTTCCTTGGGGGTGAGCCCAAGGTGCACAAATACAGATATTTCTTATTC 136
1279 AGATTCCTGGACACGTGCACCAATTATTAACACCCCACTTACGCCCAATCACTGGG 1338
137 AGATTCCTGGACACGTGCACCAATTATTAACACCCCACTTACGCCCAATCACTGGG 196
1339 AGGAAGTGAATCTCTCTTCTGATTCACAGAGTACTTTCACGGGTCAAGACAG 1398
197 AGGAAGTGAATCTCTCTTCTGATTCACAGAGTACTTTCACGGGTCAAGACAG 255
1399 CAGCTATTATGATGAAACCTTAAAGGGCAACAATTCCTGCTTCTAGGCTTAA 1458
256 CAGCTATTATGATGAAACCTTAAAGGGCAACAATTCCTGCTTCTAGGCTTAA 315
1459 CAGGAATTGGCAACAATCTGTGGCTGTCAGCAAAAGATGTCATATTTAAGATCTT 1518
316 CAGGAATTGGCAACAATCTGTGGCTGTCAGCAAAAGATGTCATATTTAAGATCTT 375
1519 GTCTTGGGCTGGGTGAGGCAAGTGAATCAAGAGGTCAAGAGTTGAGACCAACT 1578
376 GTCTTGGGCTGGGTGAGGCAAGTGAATCAAGAGGTCAAGAGTTGAGACCAACT 435
1579 GGCCACATGATGAAACCCCATCTTACCAAAAAAATCAATCACTGGCCGTGG 1638
436 GGCCACATGATGAAACCCCATCTTACCAAAAAAATCAATCACTGGCCGTGG 495
1639 TGTGCTGTAGTCCCAAGCAGAGGTTGAGGGGAAATTCCTGAACCCAGAGGTGT 1698
496 TGTGCTGTAGTCCCAAGCAGAGGTTGAGGGGAAATTCCTGAACCCAGAGGTGT 555
1699 GGTTCGAGTGAATGAGCAATCCAGCTCGGGCGGAGAGTGAAGCTGTCTCAA 1758
556 GGTTCGAGTGAATGAGCAATCCAGCTCGGGCGGAGAGTGAAGCTGTCTCAA 615

QY 1759 AAAAAAAAAAAGATGCTCTCAACCTTGGCTTCTACTGCAACATTTGG 1812
Db 616 AAAAAAAAAAAGATGCTCTCAACCTTGGCTTCTACTGCAACATTTGG 669

RESULT 2
CA441116/c
LOCUS
DEFINITION
UI-H-DP0-ava-c-06-0-UI.61 NCI CGAP Fsi Homo sapiens cDNA clone
CA441116
VERSION
KEYWORDS
CA441116.1 GI:24805336
EST.
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 668)
NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Mary Hendrix
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
The following repetitive elements were found in this cDNA
sequence: 1-21, >Al_rich#low_complexity 483-667, >Alu (matched
complement)
Seq primer: M13 FORWARD
POLYA=Yes

JOURNAL
COMMENT

FEATURES

source

Location/Qualifiers
1..668
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DP0-ava-c-06-0-UI"
/tissue_type="Fibrosarcoma"
/lab_host="DH10B (life technologies)"
/clone_lib="NCI CGAP Fsi"
/note="Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Fsi is a cDNA library containing the following tissue(s): Fibrosarcoma Cell line HT-1080 (ATCC number CCL-121). The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dr)18 tail. The sequence tag for this library is GTTCTACGAG.
TAG_TISSUE=fibrosarcoma
TAG_LIB=UI-H-DP0
TAG_SEQ=GTTCTACGAG"

ORIGIN

Query Match 27.9%; Score 619.2; DB 14; Length 668;
Best Local Similarity 99.2%; Pred. No. 2.1e-136;
Matches 633; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
QY 1581 CCAACATGATGAAACCCCATCTTACCAAAAAAATAC--AAATGAGCTGGCGGCGG 1638
Db 668 CCAACATGATGAAACCCCATCTTACCAAAAAAATACAAATTAAGCTGGCGGCGG 609
QY 1639 TGTGCTGTAGTCCCAAGCAGAGGTTGAGGGGAAATTCCTGAACCCAGAGGTGT 1698

```
Db      608  TGTGCGCTGTAGTCCCAACGACGAGGAGTTGAGGGGAGATTGCTGAACCCAGGAGGTGGT 549
Qy      1699  GGTTCAGTGAAGATTGAGCAATCCAGCTGGGCGACGAGTGAAGACTGTCTCA 1758
Db      548  GGTTCAGTGAAGATTGAGCAATCCAGCTGGGCGACGAGTGAAGACTGTCTCA 489
Qy      1759  AAAAAAAAAAAGATCGTCTCAACCTTTGGCCCTCTACTGCAACATTTTGGTATTGG 1818
Db      488  AAAAAAAAAAAGATCGTCTCAACCTTTGGCCCTCTACTGCAACATTTTGGTATTGG 429
Qy      1819  AAATGAAGTACCTTCCATATCTTATGCTTTAATATCTTCACTGAAGATGAAGC 1878
Db      428  AAATGAAGTACCTTCCATATCTTATGCTTTAATATCTTCACTGAAGATGAAGC 369
Qy      1879  ACATTCCTAACCTGCTTCTTAATGAGGATGCTTGGCAGCAGAGTCTCACTGTGTGTA 1938
Db      368  ACATTCCTAACCTGCTTCTTAATGAGGATGCTTGGCAGCAGAGTCTCTCACTGTGTGTA 309
Qy      1939  CACCAGACGACATGATCCAGTCCAGCAGCATCAGCTGTCCACACTGAAGAACGTGCC 1998
Db      308  CACCAGACGACATGATCCAGTCCAGCAGCATCAGCTGTCCACACTGAAGAACGTGCC 249
Qy      1999  TACACAGCCTGATCAATATGTTAGCTTAATATGATTAATCCAGACTACTTCAAGCCT 2058
Db      248  TACACAGCCTGATCAATATGTTAGCTTAATATGATTAATCCAGACTACTTCAAGCCT 189
Qy      2059  TTAAATGCTTTTATCTTAATAAACTGTGAAGAGTGAAGTGAACATTTGAACATTTAAC 2118
Db      188  TTAAATGCTTTTATCTTAATAAACTGTGAAGAGTGAAGTGAACATTTGAACATTTAAC 129
Qy      2119  TCAGACTCTGATTCAGAGTCGAGAACCCCTTAGTCTATCTGAATCCAGACAGCCACAC 2178
Db      128  TCAGACTCTGATTCAGAGTCGAGAACCCCTTAGTCTATCTGAATCCAGACAGCCACAC 69
Qy      2179  CTTAGTATACCTGCCCAACTAATAGATTATTAATATAC 2216
Db      68  CTTAGTATACCTGCCCAACTAATAGATTATTAATATAC 31

RESULT 3
BU629816/c 654 bp mRNA linear EST 23-SEP-2002
LOCUS      BU629816
DEFINITION UI-H-FLO-bdm-e-06-0-UI-61 NCI CGAP FLO Homo sapiens cDNA clone
ACCESSION  BU629816
VERSION    BU629816.1 GI:23296239
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 654)
AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TUMOR     Tumor Gene Index
COMMENT    Unpublished (1997)
CONTACT   Robert Strausberg, Ph.D.
Tissue    Email: cgabbs-r@mail.nih.gov
CDNA      Tissue Procurement: James Martin
Library   CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
Library   DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Library   Clone Distribution: Clone distribution information can be obtained
Library   from Dr. M. Bento Soares, bento-soares@uiowa.edu
Library   The following repetitive elements were found in this cDNA
Library   sequence: 303-531. >ALU (matched complement)
Library   Seq primer: M13 FORWARD
Library   PolyA=yes.

FEATURES
source     1..654
            /organism="Homo sapiens"
            /mol_type="mRNA"
```

```
/db_xref="taxon:9606"
/clone="UI-H-FLO-bdm-e-06-0-UI"
/tissue_type="Cell lines"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FLO"
/notes="Organ: Chondrosarcoma; Vector: pT73-Pac
(Pharmacia) with a modified polylinker; Site 1: Ecor I;
Site 2: Not I; NCI CGAP FLO is a cDNA library derived from
a pool of mRNA obtained from 4 cell lines from grade III
chondrosarcoma tissues. The library was constructed
according to Bonaldi, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an Ecor I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dr)18 tail. The
sequence tag for this library is GAGTGGGTG. The cell line
was provided by Dr James Martin from University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UI-H-FLO
TAG_SEQ=GAGTGGGTG"

ORIGIN

Query Match      27.6%; Score 611.2; DB 13; Length 654;
Best Local Similarity 98.4%; Pred. No. 1.7e-134;
Matches 627; Conservative 0; Mismatches 9; Indels 1; Gaps 1;

Qy      1414  AAACCTTAAAGGGGCAACATTCATCTTCTTCTAGGCTAGACAGAACTTGGCAA 1473
Db      654  AAACCTTAAAGGGGCAACATTCATCTTCTTCTTCTAGGCTAGACAGAACTTGGCAA 555
Qy      1474  CATCTGTGCGCTGTTCAGCA-A-AGATGTTCAATATTAAGATCTTGTCTTGGCTGGGT 1532
Db      594  CATCTGTGCGCTGTTCAGCAANAGATGTTAATATTAAGATCTTGTCTTGGCTGGGT 535
Qy      1533  GTGAGAGGCAAGTCAATCAAGAGGTCAGAGATTTAGACCAACTTGGCCAACTGATGA 1592
Db      534  GTGAGAGGCAAGTCAATCAAGAGGTCAGAGATTTAGACCAACTTGGCCAACTGATGA 475
Qy      1593  AACCCCATCTTACCAAAAAAATACAATACAGCTGCGGTGCTGGTGGCTGTAGTGC 1652
Db      474  AACCCCATCTTACCAAAAAAATACAATACAGCTGCGGTGCTGGTGGCTGTAGTGC 415
Qy      1653  CAACGACAGAGGTTGAGGGGAGATTTGTTGAACCCAGAGAGTGTGTTCAGTAGAGAT 1712
Db      414  CAACGACAGAGGTTGAGGGGAGATTTGTTGAACCCAGAGAGTGTGTTCAGTAGAGAT 355
Qy      1713  TGAGCAACTGCAATCCAGCTTGGGCGACGAGTGAAGTGTCTCAAAAAAAAAAAAAA 1772
Db      354  TGAGCAACTGCAATCCAGCTTGGGCGACGAGTGAAGTGTCTCAAAAAAAAAAAAAA 295
Qy      1773  GGAATGCTCAACCTTGGCCCTCTCTACGCAACATTTGGTATTTGAATGAAGTACCT 1832
Db      294  GGAATGCTCAACCTTGGCCCTCTCTCTACGCAACATTTGGTATTTGAATGAAGTACCT 235
Qy      1833  TCCATATCTTATGCTGTTAATATCTTCAATCTCACTAGGAGTGAAGACATTTCTTAACCTG 1892
Db      234  TCCATATCTTATGCTGTTAATATCTTCAATCTCACTAGGAGTGAAGACATTTCTTAACCTG 175
Qy      1893  CTTCTTAATGGGAGTCTTGGCAGCAGAGTCTCACTGTGTGTACACAGACAGACAC 1952
Db      174  CTTCTTAATGGGAGTCTTGGCAGCAGAGTCTCACTGTGTGTACACAGACAGACAC 115
Qy      1953  TGATCAGTCAAGCATCAGCTGTCCACAGTGAAGAGTGTCTTAACAAGCTGGA 2012
Db      114  TGATCAGTCAAGCATCAGCTGTCCACAGTGAAGAGTGTCTTAACAAGCTGGA 55
Qy      2013  TCAAAATGTTAGCTTAATAGATTAATCCAGACTA 2049
Db      54  TCAAAATGTTAGCTTAATAGATTAATCCAGACTA 18
```

CA308356
LOCUS

LOCUS	669 bp	mRNA	linear	EST 01-NOV-2002
CA308356				
DEFINITION	UI-H-PT1-bhz-k-14-0-UI_s1	NCI_CGAP_P1	Homo sapiens	cDNA clone
ACCESSION	UI-H-PT1-bhz-k-14-0-UI_3'			mRNA sequence.
CA308356				

SOURCE
ORGANIZATION

ORGANISM	Phylogenetic (Venn) Analysis
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
REFERENCE	
1 (bases 1 to 669)	
NT_004482.1	

JOURNAL COMMENT
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
misc@cgap.nci.nih.gov

1
2
3
4
5
6
7

Statistical Procurement: Dr. Gary W. Hunninghake, U of I
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@iowa.edu
The following repetitive elements were found in this cDNA
sequence: 481-669, >ALU (matched complement)
Seq primer: M13 FORWARD
PolyA=Yes.

Source

```

source
1. 669 "GenBank"
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="U1-H-F11-bnz-k-14-0-U1"
/tissue_type="Alveolar Macrophage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_11b="NCI_CGAP_F11"
/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI_CGAP_F11 is a normalized cDNA library constructed from
a pool of 81 RNA samples from Alveolar Macrophages
challenged with different treatments. The library was
normalized according to Bonaldo, Lennon and Soares, Genome
Research, 6:791-806, 1996. First strand cDNA synthesis was
primed with an oligo-dT primer containing a Not I site.
Double stranded cDNA was ligated to an EcoR I adaptor,
digested with Not I, and cloned directionally into
pT7T3-Pac vector. The oligonucleotide used to prime the
sequence that is located between the Not I site and the
GGCCATCGCG. The tissue was provided by Dr. Gary W.
Hunninghake of the University of Iowa.
TAG_L1b=U1-H-F11
TAG_SEQ=GGCCATCGCG"

```

ATTORNEY

Query Match	27.6%;	Score 611.2;	DB 14;	Length 669;
Best Local Similarity	99.1%;	Pred. No. 176-134;		
Matches 655;	Conservative	0;	Mismatches 4;	Indels 2; Gaps 2
QY	1577	CTGGCCAACATATGAAACCCCACTCTCTA	-CCAAAAAAATACAAATGACGTGGCCGTGC	1633
Db	669	CTGGCCAAACATGATTAACCCCACTCTACCCAAAAAAATACAAATTAAGCTGGGCGCTGC		610
QY	1636	TGGTGCTCCTGAATGCCCAACGACGAGGTGGAGGAGAAATTGCTTGAACCCAGAGCT		1695
Db	609	TGGTGCTCCTGAATGCCCAACGACGAGGAGGTGGAGGAGAAATTGCTTGAACCCAGAGAGT		550

QY	1636	GGTGGTTCGACATGAGATTGAGCACTGCAATCCAGCCTGGGCGACGAGTAGACCTGTC	1755
Db	549	GGTGGTTCGACATGAGATTGAGCAACTGCAATCCAGCCTGGGCGACGAGTAGACCTGTC	1755
QY	1756	CAAAAAAAAAAAAAAAAAAGATCGCTCAACCTTTGCCCTCACTGCAACAATTTTGGTAT	490
Db	489	C-AAAAAAAAAAAAAAAAAGATCGCTCAACCTTTGCCCTCACTGCAACAATTTTGGTAT	490
QY	1816	TTGAAATGAAAGTACCTCCATCTACTATGCTTTAAATCTTCACTTCACTAGAGATGA	431
Db	430	TTGAAATGAAAGTACCTCCATCTACTATGCTTTAAATCTTCACTTCACTAGAGATGA	431
QY	1876	AGACATTTCTTAACCGCTTCTCTAAATGGGAGTGTTCGCGACGAGGTCTCACTGTGT	371
Db	370	AGACATTTCTTAACCGCTTCTCTAAATGGGAGTGTTCGCGACGAGGTCTCACTGTGT	371
QY	1936	GTAACACGACGAGACATGATTCAGTCACAGCCATACAGCTGTGCCACATCGAAGAACTGT	311
Db	310	GTAACACGACGAGACATGATTCAGTCACAGCCATACAGCTGTGCCACATCGAAGAACTGT	311
QY	1996	TCTTCAACAAGCCTGGAATCAAAATGGTTAGCTTAAATATGATTAATAATCCAGACTTCTCAG	251
Db	250	TCTTCAACAAGCCTGGAATCAAAATGGTTAGCTTAAATATGATTAATAATCCAGACTTCTCAG	251
QY	2056	CGTTTAAAGCGTTTATTCATATAAACTGTAAAGCTAGACGAAACATTTGAAACATTT	191
Db	190	CGTTTAAAGCGTTTATTCATATAAACTGTAAAGCTAGACGAAACATTTGAAACATTT	191
QY	2116	AATCAGACTGTGATTCAAGAGTCGGGAACCTTATGTTCTATCTGAATCCAAAGACGCCA	131
Db	130	AATCAGACTGTGATTCAAGAGTCGGGAACCTTATGTTCTATCTGAATCCAAAGACGCCA	131
QY	2176	CACCTTAGTACTGCCCAACCAATAGATTAAATAATAC	71
Db	70	CACCTTAGTACTGCCCAACCAATAGATTAAATAATAC	71

BE410951

BE410951	773 bp	MRNA	linear	EST 21-JUL-2000
LOCUS				
DEFINITION	601303662F1 NIH_MGC_21 Homo sapiens cDNA			
	clone IMAGE:5638175 5'			
	MRNA sequence.			
ACCESSION				
	nm_004484			

VERSION

VERSION	BE410951.1	GI:934740
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

ORGANIZATION

REFERENCE	Eukaryota, Metazoa, Chordata, Craniata; Vertebrata; Euteleostomi 1 (Bases 1 to 773)
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
COMMENT	Contact: Robert Strauchman at 2

COMMENT

CONTACT: Robert Strassberg, Ph.D.
Email: c9abbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
Clone Distribution by: Incyte Genomics, Inc.
found through the I.M.A.G.E. Consortium information can be
placed: L16C338 row: 1 column: 16
High quality sequence stop: 662.
Location/notes:

Source

```

source
1. 773
/organism="Homo sapiens"
/mol_type="mRNA"
/d_xref="taxon:9606"
/clone="IMAGE:3638175"
/tissue_type="choriocarcinoma"
/lab_host="VDH10B (phage-resistant)"
/clone_lib="NH_MGC_21"

```


/note="Organ: placenta; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

```
Query Match      27.2%; Score 603.4; DB 10; Length 773;
Best Local Similarity 96.9%; Pred. No. 1.2e-132;
Matches 648; Conservative 0; Mismatches 16; Indels 5; Gaps 3;

QY 68 GATGCTCAGCTGGGCTTCATTGGCCGCTTCTTAGGCTTCATGATCATCATCC 127
    |||||
DB 1 GGAATGCTCAGCTAGCTATTCGCCCCCTTCTTAGGCTTCATGATCATCATCC 60

QY 128 TCCGAGGCGAACTCAAGACCAAACTAGTGGCCCAAGAGGGAATCCCTTCGACGA 187
    |||||
DB 61 TCCGAGGCGAACTCAAGACCAAACTAGTGGCCCAAGAGGGAATCCCTTCGACGA 120

QY 188 GGGCTGCCCC-AAAAACCAAGAGGCGACCAAGAGGAGGAGGAGGAGGAGGAGG 246
    |||||
DB 121 GGGCTGCCCCAAAAAACCAAGAGGCGACCAAGAGGAGGAGGAGGAGGAGGAGG 180

QY 247 AGGCTTGGAAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGAGCCAG 306
    |||||
DB 181 AGGCTTGGAAAGCTTAAAGCTGTGAGCGCTTCAAGTGTGCCCACTGTATCAGAGCCAG 240

QY 307 GCTACTACAGTGGCCCCAGAGCGCCCTTCAGCCCACTCCATGTTCTTCCCTTCAGAAC 366
    |||||
DB 241 GCTACTACAGTGGCCCCAGAGCGCCCTTCAGCCCACTCCATGTTCTTCCCTTCAGAAC 300

QY 367 CATCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGAGATAGACCAAGAGGAGGAGGAGG 426
    |||||
DB 301 CATCAGGCGCGTCAAAAGCTTCAAGTGTCAAGGAGATAGACCAAGAGGAGGAGGAGG 360

QY 427 AGACTGTAGTTCGGGGCCCAAGAAAGTTTGAATTCCTCAGAGAGGAGTGGGGCTT 486
    |||||
DB 361 AGACTGTAGTTCGGGGCCCAAGAAAGTTTGAATTCCTCAGAGAGGAGTGGGGCTT 420

QY 487 TGAATGAGGACCAAGAGATCTCAAGTGAAGAGGAGAAAGTGGAGTTTAACCTAGCG 546
    |||||
DB 421 TGAATGAGGACCAAGAGATCTCAAGTGAAGAGGAGAAAGTGGAGTTTAACCTAGCG 480

QY 547 ATATGCGAGAGATCCCGGAAATCACTCAAGAACCTTTGAAACATCAACCAACACA 606
    |||||
DB 481 ATATGCGAGAGATCCCGGAAATCACTCAAGAACCTTTGAAACATCAACCAACACA 540

QY 607 TACACACTACATCAAGATCAATGATCTTATTTGGGCTTTGAAAAAGAGTCTGTCC 666
    |||||
DB 541 TACACACTACATCAAGATCAATGATCTTATTTGGGCTTTGAAAAAGAGTCTGTCC 600

QY 667 TCACACCTGAACCGGGGCACTGATTCGCTGGCCCAACCCAGCTTCCCTTCTCTGA 726
    |||||
DB 601 TCACACCTGAACCGGGGCACTGATTCGCTGGCCCAACCCAGCTTCCCTTCTCTGA 656

QY 727 GCCTACCTT 735
    |||||
DB 657 GCCTACCTT 665

RESULT 6
CA397980/c 629 bp mRNA linear EST 06-NOV-2002
LOCUS CA397980.1 Human Retinal pigment epithelium/choroid cDNA
DEFINITION (Un-normalized, unambiguated): cs Homo sapiens cDNA clone c898c05
3', mRNA sequence.
ACCESSION CA397980
VERSION CA397980.1 GI:24735787
KEYWORDS EST.
SOURCE Homo sapiens (human)
```

ORGANISM

Homo sapiens

REFERENCE

Bukacinska, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 629)

AUTHORS

Wistow, G., Bernstein, S.T., Wyatt, M.K., Farris, R.N., Behal, A.,

TITLE

Touchman, J.W., Bouffard, G., Smith, D., and Peterson, K.

JOURNAL

Expressed sequence tag analysis of human RPE/choroid for the

MEDLINE

and splice variants

COMMENT

Section on Molecular Structure and Function

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

FEATURES

Location/Qualifiers

source

1..629

Db 329 TTTCACCCCAATTATTAACACCCCACTTACGCCCAATCAGTGGAGAGTGAATCTTC 270
QY 1354 CCTTTCTGGATTCTCAGCAGTTTCTTCAAGGGTCAGAAACAGCAGATTAATGATG 1413
Db 269 CCTTTCTGGATTCTCAGCAGTTTCTTCAAGGGTCAGAAACAGCAGATTAATGATG 211
QY 1414 AACCTTAAAGGCAACAATTCATCTTCTTCAAGGCTAAGACAGAACTTGGCAA 1473
Db 210 AAAACTTAAAGGCAACAATTCATCTTCTTCAAGGCTAAGACAGAACTTGGCAA 151
QY 1474 CATCTGTGGCCCTGTTCAGCAAGATGTTCAATTTAAGAACTTCTTCTTGGGCTGGGTG 1533
Db 150 CATCTGTGGCCCTGTTCAGCAAGATGTTAATTTAAGAACTTCTTCTTGGGCTGGGTG 91
QY 1534 TGGAGGCAAGTGAATCAGAGAGTTCAGAGTTTGAACCAACCTGGCCCAATGATGAA 1593
Db 90 TGGAGGCAAGTGAATCAGAGAGTTCAGAGTTTGAACCAACCTGGCCCAATGATGAA 31
QY 1594 ACCCAATCTCTACCAAAAAAATTAACAA 1621
Db 30 ACCCAATCTCTACCAAAAAAATTAACAA 3

RESULT 7
Bg950642 662 bp mRNA linear EST 12-JUN-2001
LOCUS Bg950642
DEFINITION MRL-CT0735-080101-001-408 CT0735 Homo sapiens cDNA, mRNA sequence.
VERSION Bg950642.1 GI:14368813
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 662)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Gouldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MRLct2-MRL-CT0735-
080101-001-408&t3=2001-01-08&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 593.
Location/Qualifiers
1..662

FEATURES
Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0735"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low

ORIGIN

stringency conditions."

Query Match 26.2%; Score 581.2; DB 12; Length 662;
Best Local Similarity 96.8%; Pred. No. 2,3e-127;
Matches 603; Conservative 0; Mismatches 19; Indels 1; Gaps 1;

QY 1152 GGTGAGTGTGAGGGGTGTCAGGCGCCCTTACGTATTTCTCACTGCTGGAACTCAAC 1211
Db 11 GGTGAGTGTGAGGGGTGTCAGGCGCCCTTACGTATTTCTCACTGCTGGAACTCAAC 70
QY 1212 AAAATACCTTCTGCTTCTTCTTGGGTCAGCCCAAGCTGTCACAAAATGATATTTCCCT 1271
Db 71 AAAATACCTTCTGCTTCTTCTTGGGTCAGCCCAAGCTGTCACAAAATGATATTTCCCT 130
QY 1272 TTATTCAGATTTCTGTCAGACCTGTACACCAATTTTAAACCCCACTTCAGCCCAATC 1331
Db 131 TTATTCAGATTTCTGTCAGACCTGTACACCAATTTTAAACCCCACTTCAGCCCAATC 190
QY 1332 ACCGAGGAGGAGATGTAATCTTCCCTTTCTGATTTCTCAAGCATTTCTTCAAGGCTCA 1391
Db 191 ACCGAGGAGGAGATGTAATCTTCCCTTTCTGATTTCTCAAGCATTTCTTCAAGGCTCA 249
QY 1392 GAAACAGCAGCTATTATGATGAAACCTTAAAGGCAACAATTTCTTCTTCTAG 1451
Db 250 GAAACAGCAGCTATTATGATGAAACCTTAAAGGCAACAATTTCTTCTTCTAG 309
QY 1452 GCTAAGACAGAACTTGGCAACATCTGTGGCCCTTTCAGCAAGAGATGTTCAATTTAA 1511
Db 310 GCTAAGACAGAACTTGGCAACATCTGTGGCCCTTTCAGCAAGAGATGTTCAATTTAA 369
QY 1512 GAATCTTGTCTTGTGGCTGGGTGTGAGAGCAAGGAAATCAGAGGTCAGAGTTTGA 1571
Db 370 GAATCTTGTCTTGTGGCTGGGTGTGAGAGCAAGGAAATCAGAGGTCAGAGTTTGA 429
QY 1572 CCAACCTGGCCCAATGATGAAACCCCACTCTTACCAAAAAAATAAATCAATGAGTGGCC 1631
Db 430 CCAACCTGGCCCAATGATGAAACCCCACTCTTACCAAAAAAATAAATCAATGAGTGGCC 489
QY 1632 GTCTGTGTGCTGCTGTGATGCCAAGCAGAGAGTTGAGGGAGAAATGCTTGAACCCAG 1691
Db 490 GTCTGTGTGCTGCTGTGATGCCAAGCAGAGAGTTGAGGGAGAAATGCTTGAACCCAG 549
QY 1692 AGGTGTGTGTGAGAGATGATGACCACTGCAATTCAGCCTGGCCGACGAGTGAACCT 1751
Db 550 AGGTGTGTGTGAGAGATGATGACCACTGCAATTCAGCCTGGCCGACGAGTGAACCT 609
QY 1752 GTTCAAAAAAATAAATAAAG 1774
Db 610 GTTCAAAAAAATAAATAAAG 632

RESULT 8
BE385296

LOCUS BE385296 593 bp mRNA linear EST 21-JUL-2000
DEFINITION MRL-CT0735-080101-001-408 CT0735 Homo sapiens cDNA clone IMAGE:361865 5',
1 (bases 1 to 593)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Gouldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MRLct2-MRL-CT0735-
080101-001-408&t3=2001-01-08&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 593.
Location/Qualifiers
1..593

/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="CT0735"
/note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No.
196,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 593)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Brites, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Gouldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov
plate: LCM287 row: 0 column: 18
High quality sequence stop: 593.

FEATURES

Source

Location/Qualifiers

1. 593
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:361865"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_1lb="NIH MGC 20"
/notes="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 26.2%; Score 581; DB 10; Length 553;
Best Local Similarity 99.8%; Pred. No. 2,6e-127;
Matches 592; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 83 CATCATTTGGCGCTTCTCTAGAGGCTGACATGATACATCTCTCCAGGGCAATC 142
DB 1 CATCATTTGGCGCTTCTCTAGAGGCTGACATGATACATCTCTCCAGGGCAATC 60
QY 143 AAGAGCAAACTCTGTCGAGGCAAGAGGAACTCTTCCAGAGAGGCTGCCCC-AAA 201
DB 61 AAGAGCAAACTCTGTCGAGGCAAGAGGAACTCTTCCAGAGAGGCTGCCCCAAA 120
QY 202 ACCACAAGGAGGCAAAACAGAACTTTAGGGGCAAGAGCAAAAGGCTTGAAGCTTA 261
DB 121 ACCACAAGGAGGCAAAACAGAACTTTAGGGGCAAGAGCAAAAGGCTTGAAGCTTA 180
QY 262 AGGCTGTGAGCGCTTCAAGTCTGCCCCAGTATCAGAGGCAAGGCTTACAGTCCCC 321
DB 181 AGGCTGTGAGCGCTTCAAGTCTGCCCCAGTATCAGAGGCAAGGCTTACAGTCCCC 240
QY 322 CAAAGAGCGCCCTCAGCGCCCACTCCATGTTTCCCTTGAAGCACTACGCGCCGTC 381
DB 241 CAAAGAGCGCCCTCAGCGCCCACTCCATGTTTCCCTTGAAGCACTACGCGCCGTC 300
QY 382 AGCTTCAAGTGTCAAGGATAGACCAAGCAAAAGCTTAAAGATCTGTGAGTTCTG 441
DB 301 AGCTTCAAGTGTCAAGGATAGACCAAGCAAAAGCTTAAAGATCTGTGAGTTCTG 360
QY 442 GGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGGCTTGAATGAGACCCAG 501
DB 361 GGGCCAAAGAAAGTTTGAATGCTCTCAGAGAGCGATGGGGCTTGAATGAGACCCAG 420
QY 502 AAGTATCTCAAGTGAAGAGAGAACTGTGAGATTAACTCAGAGATATGCCAGAGATCC 561
DB 421 AAGTATCTCAAGTGAAGAGAGAACTGTGAGATTAACTCAGAGATATGCCAGAGATCC 480
QY 562 CCGAAATACCTCAAGAAAGCTTTGGAAACATCAACCAACCAATACACATCACTCACTCA 621
DB 481 CCGAAATACCTCAAGAAAGCTTTGGAAACATCAACCAACCAATACACATCACTCACTCA 540
QY 622 AAGATCAATGATCTTATTTGGGCTTGGAAAAAGGTCTGTCTTCAACTG 674
DB 541 AAGATCAATGATCTTATTTGGGCTTGGAAAAAGGTCTGTCTTCAACTG 593

RESULT 9
CA397981 CA397981 636 bp mRNA linear EST 06-NOV-2002
LOCUS
DEFINITION C098C05.y1 Human Retinal pigment epithelium/choroid cDNA

(Un-normalized, unamplified): cs Homo sapiens cDNA clone C098C05
5', mRNA sequence.

ACCESSION
VERSION
CA397981.1 GI:24735789
EST.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens

REFERENCE

AUTHORS

TITLE

Wistow, G., Bernstein, S.L., Wyatt, M.K., Farria, R.N., Behal, A.,
Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.
Expressed sequence tag analysis of human RPE/choroid for the
NIH Bank Project: Over 6000 non-redundant transcripts, novel genes
and splice variants
Mol. Vis. 8 (4), 205-220 (2002)

JOURNAL

MEDLINE

PUBMED

CONTACT: Wistow G
Section on Molecular Structure and Function
National Eye Institute
6/331, NIH, Bethesda, MD 20892-2740, USA
Tel: 301 402 3452
Fax: 301 496 0078
Email: graeme@helix.nih.gov
plate: 98 row: C column: 05
Seq primer: M13RP1 reverse primer (ABI).

FEATURES

Source

Location/Qualifiers

1. 636
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="C098C05"
/tissue_type="RPE/choroid"
/dev_stage="Adult"
/lab_host="EMDH10B"
/clone_1lb="Human Retinal pigment epithelium/choroid cDNA
(un-normalized, unamplified): cs"
/notes="Organ: Eye; Vector: pCMVSPORT6; Two different donor
eyes (75-80 years old) yielded approximately 600 mg of
dissected RPE/choroid tissue. This in turn yielded 340 ug
of total RNA and 7 ug of mRNA. A directionally cloned cDNA
library in the pCMVSPORT6 vector was constructed at Life
Technologies (Rockville, MD; now part of Invitrogen Corp),
essentially following the protocols of the Superscript
Plasmid System (Invitrogen Corp).
<http://www.invitrogen.com/>. The library code
designation was cs. For this library, cDNA inserts were
cloned into the NotI/MluI sites of the vector. EST
analysis was performed on the unamplified library at the
NIH Intramural Sequencing Center (NISC)."

ORIGIN

Query Match 25.1%; Score 555.8; DB 14; Length 636;
Best Local Similarity 99.6%; Pred. No. 2,6e-121;
Matches 557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 17 CAGCTGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGCAGAGAGATCTCA 76
DB 78 CAGCTGAACAAAGAGAGATGAGTTCCAGCCCAATCAGAGAGCAGAGAGATCTCA 137
QY 77 CGGTGCATCATTTGGCGCTTCTCTAGAGGCTGACATGATACATCTCTCCAGGGC 136
DB 138 CGGTGCATCATTTGGCGCTTCTCTAGAGGCTGACATGATACATCTCTCCAGGGC 197
QY 137 AAAGTCAAGAGCAAACTACTGTGCGCCCAAGAGGAACTCTTCTCAAGAGGCTGCG 196
DB 198 AAAGTCAAGAGCAAACTACTGTGCGCCCAAGAGGAACTCTTCTCAAGAGGCTGCG 257
QY 197 CAAAAACCAAGAGGAGCCCAACAGAACTTAAAGGGCCAGAGAGCAACAGGCTTGAA 256
DB 258 CAAAAACCAAGAGGAGCCCAACAGAACTTAAAGGGCCAGAGAGCAACAGGCTTGAA 317

QY 257 GCTTAAGGCTGNGAGCGCTTCAAGTCTTCCCTCACTGATCATGAGGCCAGGCTACTACAG 316
 DB 318 GTTTAAGGCTGNGAGCGCTTCAAGTCTTCCCTCACTGATCATGAGGCCAGGCTACTACAG 377
 QY 317 TGCCCAAGAGAGCGCCCTCAGGCCCTCCATGTTCTTCCCTCAGAAACCATGAGGCC 376
 DB 378 TGCCCAAGAGAGCGCCCTCAGGCCCTCCATGTTCTTCCCTCAGAAACCATGAGGCC 437
 QY 377 GTCAAGGCTTCAAGTCTCAGAGGATGACACCAAGACAAAGCTTAAAGACTGTGAG 436
 DB 438 GTCAAGGCTTCAAGTCTCAGAGGATGACACCAAGACAAAGCTTAAAGACTGTGAG 497
 QY 437 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGGCGATGGGGCTTGAATGAGCA 496
 DB 498 TTCTGGGGCCAAAGAAAGTTTGAATGCTCTCAGAGGCGATGGGGCTTGAATGAGCA 557
 QY 497 CCCAGAGATATCTCAAGTGAAGAGAGAAACTGTGAGATTAACTGACGATATGCCAGA 556
 DB 558 CCCAGAGATATCTCAAGTGAAGAGAGAAACTGTGAGATTAACTGACGATATGCCAGA 617
 QY 557 GATCCCGGAAATACCTTC 575
 DB 618 GATCCCGGAAATACCTTC 636

RESULT 10
 CA395098 545 bp mRNA linear EST 06-NOV-2002
 CA395098
 LOCUS
 DEFINITION
 (Un-normalized, unamplified): cs Homo sapiens cDNA clone ca60c05
 5', mRNA sequence.
 CA395098
 CA395098.1 GI:24730259
 EST.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 545)
 Wistow, G., Bernstein, S.L., Wyatt, M.K., Parrie, R.N., Behal, A.,
 Touchman, J.W., Bouffard, G., Smith, D. and Peterson, K.
 Expressed sequence tag analysis of human RPE/choroid for the
 NEIBank Project: Over 6000 non-redundant transcripts, novel genes
 and splice variants
 Mol. Vis. 8 (4), 205-220 (2002)

JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 Contact: Wistow G
 Section on Molecular Structure and Function
 National Eye Institute
 6/331, NIH, Bethesda, MD 20892-2740, USA
 Tel: 301 402 3452
 Fax: 301 496 0078
 Email: gyaeme@helix.nih.gov
 Plate: 60 row: c column: 05
 Seq primer: M13RPI reverse primer (ABI).
 Location/Qualifiers

FEATURES
 source

1. 545
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="ca60c05"
 /tissue_type="RPE/choroid"
 /dev_stage="Adult"
 /lab_host="EMDH1.0B"
 /clone_lib="Human Retinal pigment epithelium/choroid cDNA
 (Un-normalized, unamplified): cs
 /note="Organ: Eye; Vector: pCMVSPORT6; Two different donor
 eyes (75-80 years old) yielded approximately 600 mg of
 dissected RPE/choroid tissue. This in turn yielded 340 ug
 of total RNA and 7 ug of mRNA. A directionally cloned cDNA
 library in the pCMVSPORT6 vector was constructed at Life
 Technologies (Rockville, MD; now part of Invitrogen Corp),

ORIGIN
 Query Match 24.6%; Score 545; DB 14; Length 545;
 Best Local Similarity 100.0%; Pred. No. 9.5e-119; Mismatches 0;
 Matches 545; Conservative 0; Indels 0; Gaps 0;
 essentially following the protocols of the SuperScript
 Plasmid System (Invitrogen Corp.
 <http://www.invitrogen.com/>). The library code
 designation was cs. For this library, cDNA inserts were
 cloned into the NotI/Mui sites of the vector. EST
 analysis was performed on the unamplified library at the
 NIH Intramural Sequencing Center (NISC)."

QY 167 GAGGGAATCCCTTCTCCAGAGGCGCTGCCAAAAACCAAGAGCCCAAGAGCGT 226
 DB 1 GAGGGAATCCCTTCTCCAGAGGCGCTGCCAAAAACCAAGAGCGCAAGAGCGT 60
 QY 227 TAGGGGCGAGGAAG 286
 DB 61 TAGGGGCGAGGAAG 120
 QY 287 CCCACTGTATCAGAGGCGAGGCTACTACAGTGCCTCCACAGAGGCCCTCAGCCCTCC 346
 DB 121 CCCACTGTATCAGAGGCGAGGCTACTACAGTGCCTCCACAGAGGCCCTCAGCCCTCC 180
 QY 347 CATGTTCTTCCCTTGAACCATCAGCGCGCTCAAGGCTTCAAGTGTACAGCATAGA 406
 DB 181 CATGTTCTTCCCTTGAACCATCAGCGCGCTCAAGGCTTCAAGTGTACAGCATAGA 240
 QY 407 CACCAAGACAAAGAGCTTAAAGAGTGTAGTCTGGGGCGAGAAAGTTTGAATGCT 466
 DB 241 CACCAAGACAAAGAGCTTAAAGAGTGTAGTCTGGGGCGAGAAAGTTTGAATGCT 300
 QY 467 CTCAGAGCGATGAGGCGCTTGAATGAGAGACCCAGAGATATCTCAAGTGAAGAGAAAC 526
 DB 301 CTCAGAGCGATGAGGCGCTTGAATGAGAGACCCAGAGATATCTCAAGTGAAGAGAAAC 360
 QY 527 TGTGAGGTTTAACTGAGAGATATGAGAGATATGAGAGATATGAGAGATATGAGAGAT 586
 DB 361 TGTGAGGTTTAACTGAGAGATATGAGAGATATGAGAGATATGAGAGATATGAGAGAT 420
 QY 587 GGAAGCAATCACCAACCAATACATACATACATACATACATACATACATACATACAT 646
 DB 421 GGAAGCAATCACCAACCAATACATACATACATACATACATACATACATACATACAT 480
 QY 647 CTTGAAAAACAGGCTGTCTCTCACCTGAACAGGGGCACTGATTCCTGTGCCCCAC 706
 DB 481 CTTGAAAAACAGGCTGTCTCTCACCTGAACAGGGGCACTGATTCCTGTGCCCCAC 540
 QY 707 CCCAG 711
 DB 541 CCCAG 545

RESULT 11
 BU657324 545 bp mRNA linear EST 30-SEP-2002
 BU657324
 LOCUS
 DEFINITION
 c122h04.21 Hembase, Erythroid Precursor Cells (LCB:cl library) Homo
 sapiens cDNA clone c122h04 5', mRNA sequence.
 BU657324
 BU657324.1 GI:23369506
 EST.
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 545)
 Gubin, A.N., Lee, Y.T., Bouffard, G.G. and Miller, J.L.
 Gene Expression in Human Erythroid Precursor Cells
 Unpublished (2002)
 Contact: Jeffery L. Miller
 Laboratory of Chemical Biology
 National Institute of Diabetes and Digestive and Kidney Diseases

Best Local Similarity 99.4%; Pred. No. 1.9e-108;
Matches 514; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

```
QY 148 CCAACTACTGTGGCCCAAG-AGGAAATCCCTTCTCCAGAGAGGCTGCCCAAAACAC 206
Db 1 CCAACTACTGTGGCCCAAGAGAAATCCCTTCTCCAGAGAGGCTGCCCAAAACAC 60
QY 207 AAGGAGCCCAACAGACGTTAGGGGCGCAGAAAGCAACAGAGCTTGAAGCTTAAGGCT 266
Db 61 AAGGAGCCCAACAGACGTTAGGGGCGCAGAAAGCAACAGAGCTTGAAGCTTAAGGCT 120
QY 267 GTGAGAGCCTTCAAGTCTGCCCACTGTATTCAGAGGCGCAGGCTACTTAAGTCCCAAG 326
Db 121 GTGAGAGCCTTCAAGTCTGCCCACTGTATTCAGAGGCGCAGGCTACTTAAGTCCCAAG 180
QY 327 AGCCCTCAGAGCCCACTCCCAATGTTCTTCCCTTGAAGACATCAGCCCGCCCAAGCTT 386
Db 181 AGCCCTCAGAGCCCACTCCCAATGTTCTTCCCTTGAAGACATCAGCCCGCCCAAGCTT 240
QY 387 CACAGTGTACAGGCAATAGACCAAAAGCAAAAGCTTAAAGCTGTGAGTTCTGGGCGC 446
Db 241 CACAGTGTACAGGCAATAGACCAAAAGCAAAAGCTTAAAGCTGTGAGTTCTGGGCGC 300
QY 447 AAGAAAGTTTGAATGCTCTCAGAGAGCCATGGGCGCTTGAAGAGCACCAGAAAGTA 506
Db 301 AAGAAAGTTTGAATGCTCTCAGAGAGCCATGGGCGCTTGAAGAGCACCAGAAAGTA 360
QY 507 TCTCAAGTGAAGAGGAAACTGTGAGTTAACTGACGATATGCCAGAGATCCCGGAA 566
Db 361 TCTCAAGTGAAGAGGAAACTGTGAGTTAACTGACGATATGCCAGAGATCCCGGAA 420
QY 567 AATCACTCTCAAAAGAACTTTGGAAATCAATCAACCAACATACACACTACCTCAAGAT 626
Db 421 AATCACTCTCAAAAGAACTTTGGAAATCAATCAACCAACATACACACTACCTCAAGAT 480
QY 627 CACATGATCCTTATTTGGGCGCTTGGAAACAGGCTGT 663
Db 481 CACATGATCCTTATTTGGGCGCTTGGAAACAGGATG 517
```

RESULT 13
LOCUS BM697639 523 bp mRNA linear EST 28-FEB-2002
DEFINITION UI-E-DX0-agn-n-07-0-UI r1 UI-E-DX0 Homo sapiens cDNA clone
ACCESSION BM697639
VERSION BM697639.1 GI:19010897
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 523)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MBRRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA

FEATURES
source
sequence: 9-237, >ALU
Seq primer: M13 Reverse.
Location/Qualifiers
1..523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DX0-agn-n-07-0-UI"
/tissue_type="fetal eyes"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site: 1: EcoR I; Site 2: Not I;
UI-E-DX0 is a cDNA library containing the following
tissue(s): fetal eyes. The library was constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT73-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is AGAATCAAGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 22.4%; Score 496.6; DB 12; Length 523;
Best Local Similarity 98.8%; Pred. No. 3.2e-107;
Matches 510; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
QY 1536 GAGGCAAGTGAATCAAGAGAGTCAAGAGTTGAGACCAACTGGCCAAATGATGAAC 1595
Db 9 GAGGCAAGTGAATCAAGAGAGTCAAGAGTTGAGACCAACTGGCCAAATGATGAAC 68
QY 1596 CCCATCTCAACAAAAAATACAAATACAGTGGCGCGCGTGTGTGCTTACTATCCCA 1655
Db 69 CCCATCTCAACAAAAAATACAAATACAGTGGCGCGCGTGTGTGCTTACTATCCCA 128
QY 1656 GCGAGAGGTTAGGGGAAATGCTTGAACCCAGAGGTGTGCTGAGTGAATGA 1715
Db 129 GCGAGAGGTTAGGGGAAATGCTTGAACCCAGAGGTGTGCTGAGTGAATGA 188
QY 1716 GCAACTGAATCCAGCCCTGGCGGACGAGTGAAGCTGTCTTAAAAAAAAGGA 1775
Db 189 GCAACTGAATCCAGCCCTGGCGGACGAGTGAAGCTGTCTTAAAAAAAAGGA 247
QY 1776 TCGTCTCAACCTTTGCCCTCTACTGCAACATTTGGATTTGAATGAAGTACCTTC 1835
Db 248 TCGTCTCAACCTTTGCCCTCTACTGCAACATTTGGATTTGAATGAAGTACCTTC 307
QY 1836 ATAATTATGCTTAATTAATCTTCACTTCACTAGAGTGAAGCAATCTTCACTGCTT 1895
Db 308 ATAATTATGCTTAATTAATCTTCACTTCACTAGAGTGAAGCAATCTTCACTGCTT 367
QY 1896 CCTAATGGGAGATGCTTGGCCAGCAGGTCTCTACCTGTGTATACACAGAGACACTGA 1955
Db 368 CCTAATGGGAGATGCTTGGCCAGCAGGTCTCTACCTGTGTATACACAGAGACACTGA 427
QY 1956 TCCAGTCAAGCCATACAGCTGTCCACACTGAAGAGTGTCTTACAAACAGCTGAATCA 2015
Db 428 TCCAGTCAAGCCATACAGCTGTCCACACTGAAGAGTGTCTTACAAACAGCTGAATCA 487
QY 2016 AATGTTAGCTTAATTAATTAATTAATCCAGACTACT 2051
Db 488 AATGTTAGCTTAATTAATTAATTAATCCAGACTACT 523
```

RESULT 14
LOCUS BM707813 529 bp mRNA linear EST 28-FEB-2002

DEFINITION UI-E-CII-afa-1-23-0-UI.r1 UI-E-CII Homo sapiens cDNA clone
ACCESSION UI-E-CII-afa-1-23-0-UI 5', mRNA sequence.
VERSION BM707813
KEYWORDS BM707813.1 GI:19021071
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 529)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: Bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.regen.com).
The following repetitive elements were found in this cDNA
sequence: 435-519, >ALU
Seq primer: M13 Reverse.
FEATURES
source
Location/Qualifiers
1..529
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-CII-afa-1-23-0-UI"
/tissue="RPE and Choroid"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/notes="Organ: eye; Vector: pRTT3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-CII is a normalized cDNA library containing the
following tissue(s): RPE and Choroid. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pRTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is ACCTA.
This library was created for the program, Gene Discovery
in the Visual System, supported by National Eye Institute
(NEI)."
ORIGIN
Query Match 22.2%; Score 492.6; DB 12; Length 529;
Best Local Similarity 97.7%; Pred. No. 2.8e-106;
Matches 509; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 1101 CACTACAGAAAGGATGCGAGAACTGCTCACTCTAGAGACTGTAGTGGTGAAGTT 1160
DB 1 CACTACAGAAAGGATGCGAGAACTGCTCACTCTAGAGACTGTAGTGGTGAAGTT 60
QY 1161 GAGGGGTCTCAGGCGCCCTTATGCTATTTTCTCACTGCTGGAACTCACCACCAAAATACCT 1220
DB 61 GAGGGGTCTCAGGCGCCCTTATGCTATTTTCTCACTGCTGGAACTCACCACCAAAATACCT 120
QY 1221 CTGCTTCTCTGGGGTCAAGCCAAAGCTGTACAAAATACGATATTTCCCTTATTCGAG 1280

DB 121 CTGCTTCTCTGGGGTCAAGCCAAAGCTGTACAAAATACGATATTTCCCTTATTCGAG 180
QY 1281 ATTCTCTGACACTGTACACCAATTTATTAACACCCCACTTACGCCCAATCAGTGGAG 1340
DB 181 ATTCTCTGACACTGTACACCAATTTATTAACACCCCACTTACGCCCAATCAGTGGAG 240
QY 1341 GAAGTGAATCTCCCTTTTCTGATTTCTCAAGCAAGTACTTTCAAGGCTGAGAACAGCA 1400
DB 241 GAAGTGAATCTCCCTTTTCTGATTTCTCAAGCAAGTACTTTCAAGGCTGAGAACAGCA 299
QY 1401 GCTATTATGATTTGAAACCTTTAAAGGCAACAAATTTCAATCTTCTTGAAGTAAAGCA 1460
DB 300 GCTATTATGATTTGAAACCTTTAAAGGCAACAAATTTCAAGTCTTCTTGAAGTAAAGCA 359
QY 1461 GGAATCTTGCAAAACATCTGTGGCTGTTCAAGCAAGATTTCAATTTTAAGATCTTGT 1520
DB 360 GGAATCTTGCAAAACATCTGTGGCTGTTCAAGCAAGATTTTAATTTTAAGATCTTGT 419
QY 1521 CTGGGCTGGGTGGAGGCAAGTAAATCAAGAGGTGAGAGTTTGAAGCAACCTGG 1580
DB 420 CTGGGCTGGGTGGAGGCAAGTAAATCAAGAGGTGAGAGTTTGAAGCAACCTGG 479
QY 1581 CCAACATGATGAAACCCCATCTCTACCAAAAAAATAACAA 1621
DB 480 CCAACATGATGAAACCCCATCTCTACCAAAAAAATAACAA 520
RESULT 15
AA307119
LOCUS
DEFINITION EST178031 Colon carcinoma (HCC) cell line Homo sapiens cDNA 5' end,
mRNA sequence.
AA307119 469 bp mRNA linear EST 18-Apr-1997
VERSION AA307119
KEYWORDS AA307119.1 GI:1959449
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryote: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 469)
AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A.,
Bull,C.J., Lee,N.H., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D.,
White,O., Sutton,G., Blake,J.A., Brandon,R.C., Man-ai,C.,
Clayton,R.A., Cline,T.R., Cotton,M.D., Batle-Hughes,J., Fine,L.D.,
Fitzgerald,L.M., Fitzhugh,W.M., Fritchman,J., Geoghegan,N.S.,
Glodek,A., Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinkle,P.S., Jr.,
Kelley,J.M., Kelley,J.C., Liu,L.-I., Marmaro,S.M., Merrick,J.M.,
Moreno-Palauques,R.F., McDonald,L.A., Nguyen,D.T., Pelligrino,S.M.,
Phillips,C.A., Ryder,S.E., Scott,J.L., Sauder,D.M., Shirley,R.,
Small,K.V., Spriggs,T.A., Uterback,T.R., Weidman,J.F., Li,Y.,
Bedarik,D.P., Cao,L., Cepeda,M.A., Coleman,T.A., Collins,B.Y.,
Dimke,D., Feng,D.-F., Ferris,A., Fischer,C., Hastings,G.A.,
He,W.W., Hu,J.S., Greene,J.M., Gruber,J., Hudson,P., Kim,A.K.,
Kozak,D.L., Kunsch,C., Hungjun,J., Li,H., Weissner,P.S., Olsen,H.,
Raymond,L., Wei,Y.F., Wang,J., Xu,C., Yu,G.L., Ruben,S.M.,
Dillon,P.D., Pannon,M.R., Rosen,C.A., Haseltine,W.A., Fields,C.,
Frazer,C.M. and Venter,J.C.
Initial assessment of human gene diversity and expression patterns
based upon 83 million nucleotides of cDNA sequence
Nature 377 (6547 Suppl), 3-174 (1995)
JOURNAL 96026280
MEDLINE 7566098
PUBMED
COMMENT Other_ESTs: THC123650
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression
information related to this EST, please check the TIGR Human Gene

Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.
Location/Qualifiers

FEATURES
source

1. 469
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):108624"
/db_xref="taxon:9606"
/tissue_type="colon"
/cell_type="KM12C"
/cell_line="KM12C(HCC)-parental human colon
carcinoma/Dukes B2"
/clone_lib="Colon carcinoma (HCC) cell line"
/note="Organ: colon; Vector: pBluescript SK-; Site_1:
EcoRI; Site_2: XhoI"

ORIGIN

Query Match 21.0%; Score 466.4; DB 9; Length 469;
Best Local Similarity 99.6%; Pred. No. 4.8e-100;
Matches 467; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 50 CAATCAGAGAGACGAGAGATGCTCAGCGCTGCATCATTTGCGCGCTTCTTAGGCTTGA 109
DB 1 CAATCAGAGAGACGAGAGATGCTCAGCGCTGCATCATTTGCGCGCTTCTTAGGCTTGA 60
QY 110 GTCCCATGATCAACATCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAG 169
DB 61 GTCCCATGATCAACATCTCCAGGGCAAACTCAAGACCAAACTACTGTGGCCCAAGAG 120
QY 170 GGAATCCCTTCTCCAGAGGGCTGCCCAAAAACCAAGGACGCAAAACGTTAG 229
DB 121 GGAATCCCTTCTCCAGAGGGCTGCCCAAAAACCAAGGACGCAAAACGTTAG 180
QY 230 GGGCCAGAGAGACAACAAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 289
DB 181 GGGCCAGAGAGACAACAAGGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGAAGCTTGA 240
QY 290 ACTGATCAGAGGCGAGGCTACTACAGTGCCCAAGAGCGCCCTCAGCCCACTCCCAT 349
DB 241 ACTGATCAGAGGCGAGGCTACTACAGTGCCCAAGAGCGCCCTCAGCCCACTCCCAT 300
QY 350 GTTCTTCCCTAGAACCATCAGAGCGCTCAAAAGCTTCAAGTCTCAGGCAATAGAC 409
DB 301 GTTCTTCCCTAGAACCATCAGAGCGCTCAAAAGCTTCAAGTCTCAGGCAATAGAC 360
QY 410 CAAAGACAAAAGCTTAAAGACTGTGAGTTCTGGGGCAAGAAAAGTTTGAATTGCTCTC 469
DB 361 CAAAGACAAAAGCTTAAAGACTGTGAGTTCTGGGGCAAGAAAAGTTTGAATTGCTCTC 420
QY 470 AGAGAGGATGGGGCTTGAATGAGCAACCAAGATATCTCAAGTGAGG 518
DB 421 AGAGAGGATGGGGCTTGAATGAGCAACCAAGATATCTCAAGTGAGG 469

Search completed: March 25, 2004, 07:21:33
Job time : 5679 secs